G3: Genes|Genomes|Genetics Early Online, published on February 15, 2019 as doi:10.1534/g3.119.0011

Chromosome-wide evolution and sex determination in the three-sexed nematode *Auanema rhodensis*

Sophie Tandonnet*, Georgios D. Koutsovoulos[†,1], Sally Adams*, Delphine Cloarec*, Manish Parihar*[,2], Mark L. Blaxter[†], Andre Pires-daSilva*

*School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK

[†]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, UK

[1]Present address: INRA, UMR1355 Institute Sophia Agrobiotech, F-06903 Sophia Antipolis, France

[2]Present address: Department of Molecular Medicine, Institute of Biotechnology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA.

* All genomic data have been submitted under Bioproject number PRJEB29492

** All RNA-seq data have been deposited in the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-7667.

Genome analysis of *Auanema rhodensis*

Auanema, genome, genetic map, Nigon elements, sex determination

Andre Pires-daSilva

School of Life Sciences, Gibbet Hill road, University of Warwick, Coventry CV4 7AL, UK

+44 2476 573 329

andre.pires@warwick.ac.uk

Abstract

Trioecy, a mating system in which males, females and hermaphrodites co-exist, is a useful system to investigate the origin and maintenance of alternative mating strategies. In the trioecious nematode *Auanema rhodensis*, males have one X chromosome (XO), whereas females and hermaphrodites have two (XX). The female *versus* hermaphrodite sex determination mechanisms have remained elusive. In this study, RNA-seq analyses show a 20% difference between the L2 hermaphrodite and female gene expression profiles. RNAi experiments targeting the DM (*doublesex*/*mab-3*) domain transcription factor *dmd-10/11* suggest that the hermaphrodite sexual fate requires the upregulation of this gene. The genetic linkage map (GLM) shows that there is chromosome-wide heterozygosity for the X chromosome in F2 hermaphrodite-derived lines originated from crosses between two parental inbred strains. These results confirm the lack of recombination of the X chromosome in hermaphrodites, as previously reported. We also describe conserved chromosome elements (Nigon elements), which have been mostly maintained throughout the evolution of Rhabditina nematodes. The seven-chromosome karyotype of *A. rhodensis*, instead of the typical six found in other rhabditine species, derives from fusion/rearrangements events involving three Nigon elements. The *A. rhodensis* X chromosome is the smallest and most polymorphic with the least proportion of conserved genes. This may reflect its atypical mode of father-to-son transmission and its lack of recombination in hermaphrodites and males. In conclusion, this study provides a framework for studying the evolution of chromosomes in rhabditine nematodes, as well as possible mechanisms for the sex determination in a three-sexed species.

Introduction

The evolution of mating systems has long interested evolutionary biologists (Maynard Smith 1978; Charnov 1982), including Darwin (Darwin 1876). The type of mating system is biologically relevant because it has various consequences for population genetics and evolution, including effective population size, the degree of homozygosity, ability to remove deleterious mutations, and rates of recombination and mutation (Glemin 2007; Wright et al. 2008; Charlesworth 2006).

The existence of mixed mating strategies, in which organisms reproduce by both self- and cross-fertilization, is a challenging problem for evolutionary biologists (Goodwillie et al. 2005; Weeks 2012). It is still a matter of controversy of whether mixed mating systems are evolutionarily stable. According to theoretical models, there are only two stable states in mating system evolution: predominant outcrossing with strong inbreeding depression or predominant selfing with weak inbreeding depression (Lande and Schemske 1985; Charlesworth et al. 1990). This led to the suggestion that mixed mating types are transitional and therefore short-lived (Lande and Schemske 1985; Charlesworth 1984). However, mixed mating systems seem to persist for long periods of time in some animal groups, even after several speciation events (Weeks et al. 2006b). The type of sex determination system (Otto et al. 1993), ecological factors and the presence of inbreeding depression in the particular species might explain the persistence of a mixed mating type (Weeks et al. 2000).

Nematodes are ideal to address these questions because they have diverse mating systems that include: parthenogenesis (Bell 1982; Triantaphyllou and Hirschmann 1964), self-compatible hermaphroditism (Maupas 1900), dioecy (males/females) (Poinar 1983; Chitwood and Chitwood 1950), androdioecy (males/hermaphrodites) (Sudhaus and Fitch 2001; Kiontke et al. 2004; Herrmann et al. 2006a) and trioecy (males, females and hermaphrodites) (Félix 2004; Maupas 1900). Many of them can be cultured in the laboratory, facilitating experimental manipulation of mating systems (Félix 2004; Stewart and Phillips 2002; Cutter 2005).

It is hypothesized that mixed systems such as androdioecy, gynodioecy (females/hermaphrodites) and trioecy are intermediate steps between dioecy and hermaphroditism in some systems (Weeks et al. 2006a). Although no gynodioecious nematodes are known, androdioecy evolved from dioecy multiple times during nematode evolution (Kiontke et al. 2004; Cho et al. 2004; Herrmann et al. 2006b; Pires-daSilva 2007). The transition from dioecy to hermaphroditism can be achieved in few steps. For instance, the down-regulation of only two genes (a sex-determining gene and a sperm-activation gene) in females of the dioecious *Caenorhabditis remanei* is sufficient to induce the development of selfing hermaphrodites (Baldi et al. 2009). Androdioecy has been described in a number of species, especially free-living, terrestrial nematodes (Maupas 1900; Sudhaus 1976; Herrmann et al. 2006a; Pires-daSilva 2007). However, it is unclear which evolutionary steps are necessary and whether this is an evolutionary stable mating strategy (Loewe and Cutter 2008; Chasnov and Chow 2002; Stewart and Phillips 2002; Cutter 2005).

By genetically manipulating the *C. elegans* sex determination it is possible to model various mating systems (Hodgkin 2002). Mutant alleles of sex determination genes have been combined to create dioecious and trioecious strains (Cutter 2005; Stewart and Phillips 2002). Dioecious strains can be indefinitely maintained by introducing a mutation (e.g., *fog-2* or *spe-27*) that prevents spermatogenesis in hermaphrodites, turning them into functional females (Schedl and Kimble 1988; Minniti et al. 1996). Synthetic trioecious populations, consisting of *C. elegans* males and females (feminised hermaphrodite mutants) mixed with wild type hermaphrodites are short-lived: males and females are rapidly outcompeted by selfing hermaphrodites (Stewart and Phillips 2002). This happens even under high mutational conditions, which are predicted to lead to a short-term advantage of obligate outcrossing over selfing (Cutter 2005). However, these trioecious populations were artificially created and it is possible that the males of *C. elegans* have lost some of their reproductive faculties. Naturally occurring trioecious nematodes, such as free-living nematodes of the genus *Auanema* (Maupas 1900; Félix 2004; Kanzaki et al. 2017) and entomopathogenic nematodes of the genus *Heterorhabditis* (Ciche 2007; Zioni (Cohen-Nissan) et al. 1992) are interesting systems to study the stability of trioecy and the mechanisms controlling the development of different sexual morphs within a same population. They also offer the opportunity to assess the costs and advantages of selfing versus outcrossing.

To investigate the possibility of trioecy being stable in specific circumstances, we have been studying the sex determination system of *Auanema* nematodes (Shakes et al. 2011; Chaudhuri et al. 2015; Chaudhuri et al. 2011). *A. rhodensis* produces males, females and hermaphrodites, both by selfing and crossing (Félix 2004). *A. rhodensis* males are XO, whereas hermaphrodites and females are XX (Shakes et al. 2011;

Chaudhuri et al. 2015; Chaudhuri et al. 2011). Due to a modified meiosis and spermatogenesis, *A. rhodensis* males produce functional spermatids with one X chromosome (haplo-X sperm), whereas sperm without the X chromosome (nullo-X sperm) are discarded (Shakes et al. 2011).

The meiosis program governing the X chromosome is also atypical in *A. rhodensis* XX hermaphrodites. We have previously shown that the X chromosome does not seem to recombine during hermaphrodite meiosis, leading to the production of mostly nullo-X oocytes and sperm with two X chromosomes (diplo-X sperm) (Tandonnet et al. 2018; Shen and Ellis 2018). Consequently, self-fertilization results mostly in XX progeny (hermaphrodites or females) (Figure 1A). Males, having only haplo-X sperm, produce exclusively male cross-progeny when fertilizing the nullo-X oocytes of hermaphrodites (Tandonnet et al. 2018) (Figure 1B). In females, the X chromosome recombines normally and thus most oocytes are haplo-X (nullo-X oocytes are also produced, although relatively rare). Cross-progeny of females with males are mostly hermaphrodites (Chaudhuri et al. 2015) (Figure 1C).

It is not clear how the hermaphrodite versus female sexual fate is determined. Selfing hermaphrodite mothers produce progressively less female progeny as they become older, whereas females produce almost exclusively hermaphrodite progeny throughout their lives (Chaudhuri et al. 2015). Hermaphrodites and females follow different developmental paths: hermaphrodites always pass through the dauer stage, whereas females (and males) do not (Félix 2004; Chaudhuri et al. 2011). Female-fated larvae that are forced to pass through the dauer stage (by the use of dauer-inducing chemicals) become adult hermaphrodites. On the other hand, blocking dauer formation

of hermaphrodite-fated larvae results in the development of adult females (Chaudhuri et al. 2011). These results indicate that dauer formation is necessary and sufficient for hermaphrodite development in *A. rhodensis* (Chaudhuri et al. 2011).

Here we sequenced the genome of *A. rhodensis* with the motivation to uncover molecular mechanisms involved in the hermaphrodite versus female fate, as well as to determine the consequences of the lack of recombination of the X chromosome in hermaphrodites and males. Previously, we determined that *A. rhodensis* has an unusual karyotype compared to most other Rhabditina nematodes, with six autosomes in addition to the X chromosome (Tandonnet et al. 2018). We show that the X chromosome is the smallest chromosome in *A. rhodensis*, is more polymorphic than autosomes and that the extra chromosome evolved from the fusion of parts of different ancestral chromosomes. Furthermore, we found that the product of the gene coding for the transcription factor *dmd-10/11* is required for the hermaphrodite fate.

Materials and Methods

Strains

We used *A. rhodensis* inbred strains APS4 and APS6 (Kanzaki et al. 2017) to produce F2-derived lines (F2Ls) and generate a genetic linkage map. Strains were maintained at 20^{o}C according to standard conditions as for *C. elegans* (Stiernagle 2006) on Nematode Growth Medium (3 g/L sodium chloride, 2.5 g/L bacto peptone, 17 g/L agar, 1 mM magnesium sulfate, 5 mg/L cholesterol, 1 mM calcium chloride, 25 mM potassium phosphate) (Brenner 1974). Plates were seeded with the *Escherichia coli* streptomycin-resistant strain OP50-1. Microbial contamination was prevented by adding 50 µg/mL of

streptomycin and 10 µg/mL of nystatin to the NGM. The inbred lines were obtained by letting a population expand from a single individual hermaphrodite (picked at the dauer stage) every few generations. The strain APS4 and APS6 underwent 50 and 11 of such rounds of bottlenecking (expansion from a single hermaphrodite), respectively.

DNA extraction, sequencing and data pre-processing

To extract nematode DNA with minimal bacterial contamination, we used the split plate method (Pires-daSilva 2013). Nematodes were cultured on one compartment of a 10 cm, two-compartment plate. The compartment with nematodes contained NGM seeded with *E. coli* OP50-1, and the second compartment contained M9 buffer. As the compartment with nematodes became crowded, dauer larvae migrated to the compartment with M9. The dauers were collected from 10 plates and washed twice with M9 buffer. The nematode pellet was stored at -80 °C. DNA was extracted and treated with RNAse using the Gentra Puregene Core A Kit (Qiagen) following the manufacturer's instructions. The DNA was dissolved in nuclease-free water for library preparation and sequencing.

The APS4 strain was chosen for genome assembly. Three independent Illumina paired-end (PE) libraries with insert sizes of 250 bp were sequenced at UT Southwestern (Dallas, Texas, USA) on an Illumina HiSeq 2500 (Table 1). Two Illumina mate-pair (MP) libraries with virtual insert sizes of ~3 kb and ~6 kb were constructed and sequenced on Illumina HiSeq 2500 at Edinburgh Genomics (University of Edinburgh, UK). The raw reads were processed to remove low-quality bases using Skewer (version 0.2.1, parameter settings "-Q 20 -l 51 -t 32") (Jiang et al. 2014). Error correction was

performed using Fiona (version 0.2.1, “-nt 48 -g 60000000”) (Schulz et al. 2014). We used Blobtools (Kumar et al. 2013) to remove microbial contamination.

To call variants in strain APS6, an Illumina paired-end library with insert size 450 bp was constructed and sequenced on Illumina HiSeq2500 at UT Southwestern (Table 1). Raw reads were preprocessed using Skewer as for the APS4 libraries (Jiang et al. 2014).

Genome assembly and annotation

De novo genome assembly was performed with SOAPdenovo2 (version r240) (Luo et al. 2012), k-mer length=71) using the PE and MP libraries for contig assembly and MP libraries for scaffolding, as this resulted in the best assembly (Table 2). The optimal k-mer length was estimated using Kmergenie (version 1.6741) (Chikhi and Medvedev 2014). We removed small contigs (<500 bp) or those having a low coverage (<5 reads/ scaffold on average). Reapr (version 1.0.17) (Hunt et al. 2013) was used to identify misassemblies within the scaffolds, and 42 questionable joins in the draft scaffolds were identified. We manually inspected these using Tablet (version 1.14.11.07) (Milne et al. 2013) and manually split 4 scaffolds that contained unjustified joins. Repeats were masked with RepeatModeler (Smit and Hubley 2008-2015) and RepeatMasker (Smit and Hubley 2013-2015).

Genome completeness was assessed using CEGMA (version 2.4) (Parra et al. 2007). Gene predictions were made using the *ab initio* and evidence-driven gene predictors GeneMark (version 2.3)(Ter-Hovhannisyan et al. 2008), SNAP (trained with CEGMA,

203 release 11/29/2013) (Korf 2004), Maker2 (version 2.31) (Campbell et al. 2014) and
204 Augustus (version 2.5)(Stanke and Waack 2003). The outputs from SNAP, GeneMark,
205 a set of *A. rhodensis* transcripts assembled by Trinity (Grabherr et al. 2011) and protein
206 similarity matches from the UniProt database were used as inputs for Maker2. We then
207 used the output from Maker2 along with hints directly generated using RNA-seq reads
208 and the set of transcripts as input for Augustus. We used the Augustus gene predictions
209 for our analyses.

210 We functionally annotated the protein coding genes by combining the results from
211 BLAST, InterProScan and Blast2GO. We performed a similarity search (-evalue 1e-5 -
212 max_target_seqs 50 -outfmt 5) against a database of all metazoan protein sequences
213 available on NCBI (28/08/2015) using BLAST+ (version 2.2.31+) (Camacho et al. 2009).
214 InterProScan (-goterms -iprlookup, version 5.14-53.0) (Jones et al. 2014) was used to
215 identify protein motifs and signatures. We used Blast2GO (version 3.1) (Götz et al.
216 2008) to integrate the InterProScan and BLAST results to add Gene Ontology (GO)
217 term annotations to *A. rhodensis* proteins. We added implicit GO terms to the existing
218 annotation using Annex (Myhre et al. 2006).

219 We identified non-coding RNA loci using Infernal (version 1.1.1) (Nawrocki and Eddy
220 2013), which uses the Rfam database. Transfer RNAs were identified using
221 RNAscanSE (version 1.3.1) (Lowe and Eddy 1997). We identified ribosomal RNAs with
222 Infernal (for 5S rRNAs) and BLASTn (BLAST+, version 2.2.31+ (Camacho et al. 2009))
223 using as a database the partial 18S (accession number EU196004.1) and 28S
224 (accession number EU195960.1) of *A. rhodensis* (Kiontke et al. 2007). We counted

unique functional RNA features using BEDTools intersect (version 2.25.0, -s -c, (Quinlan and Hall 2010)).

RAD-seq and Genetic Map Construction

A genetic map was constructed using markers obtained from restriction site-associated DNA sequencing (RAD-seq) of 95 F2-derived lines (F2Ls) originating from *A. rhodensis* inbred strains APS4 and APS6 (Kanzaki et al. 2017). To generate the F2Ls, crosses between APS4 females and APS6 males were performed to generate several F1 hermaphrodites, which were allowed to reproduce by selfing. Each progeny F2L was established from single F2 hermaphrodite progenitors, which were left to expand. For some lines, bottleneck events may have occurred after F2. We used the split plate method as described above to isolate DNA from the lines. This method relies on the isolation of dauers. Since dauers of *A. rhodensis* always develop into hermaphrodites (Chaudhuri et al. 2011; Félix 2004), the DNA isolation was derived only from this sex. Paired-end RAD-seq using PstI restriction digestion was carried out for each of the parental strains and the 95 F2Ls (Baird et al. 2008). The raw RAD-seq reads were demultiplexed and low-quality regions were removed using process_RAD_tags from the Stacks package (version 1.35) (Catchen et al. 2013). We then used the denovo_map.pl Stacks pipeline to determine the genotype of each locus (region sequenced adjacent to the PstI cut site) for each progeny sample.

The genetic map was constructed using the R packages OneMap (version 2.0-4)(Margarido et al. 2007) and r/qtl (version 1.38-4) (Broman et al. 2003). A LOD (logarithm of odds) score of 20 and a recombination fraction of 0.5 were used as parameters to arrange the loci into linkage groups. The initial genetic map was refined

by removing duplicated markers (markers with exactly the same genotype across all samples) and those with missing genotypes in 50% or more of the samples (function 'drop.markers'). Large gaps (loose markers) in the genetic map were fixed by dropping 3 markers. The Kosambi mapping function was used to determine the genetic distances between markers. However, the genetic distances could not be estimated precisely, as the level of recombination in the F2Ls is unknown.

Synteny analysis and identification of the X chromosome.

The software Chromonomer (version 1.07) (Amores et al. 2014) was used to anchor the genomic scaffolds to the genetic map, yielding a chromosomal assembly with scaffolds ordered, where possible, in each linkage group. The resulting chromosomal blocks were aligned to the *C. elegans* and *Pristionchus pacificus* genomes using PROmer (version 3.07) (Kurtz et al. 2004) with default parameters. Macro-synteny was visualized using Circos (version 0.69) (Krzywinski et al. 2009). One linkage group (LG5) aligned almost exclusively to the *C. elegans* X chromosome. We genotyped 5 polymorphic markers from this linkage group in F1 hybrid males (from an APS4 x APS6 cross) confirmed it to be the X chromosome (Tandonnet et al. 2018).

Genome analyses

Orthologous proteins between *C. elegans* (PRJNA13758.WS264), *Haemonchus contortus* (HCON_v4, early access granted by Stephen Doyle), *P. pacificus* (El Paco v1), *Oscheius tipulae* (CEW1_nOt2) and *A. rhodensis* (chromosomal assembly) were identified through reciprocal best hit BLAST searches (BLASTp, "-evalue 0.01 -max_target_seqs 100, -outfmt 6"). Localisations of orthologous proteins were visualised

using Circos plots (version 0.69) (Krzywinski et al. 2009). For *O. tipulae*, we used the correspondence of the genomic scaffolds to chromosomes (Besnard et al. 2017). For the visualisations of *O. tipulae* chromosomes, the scaffolds were collated in numerical order (the true order is currently not known). Gene (protein-coding and functional RNA) density was plotted for each chromosome using the R package karyoploteR (version 1.5.1)(Gel and Serra 2017). Protein-coding genes conserved between *A. rhodensis* and *Drosophila melanogaster* (GCF_000001215.4_Release_6_plus_ISO1_MT) were identified by performing reciprocal BLASTp searches (BLAST+, -evalue 0.01 -max_target_seqs 100, -outfmt 6). The localization of the conserved genes along the chromosomes was visualized using karyoploteR.

Variants (single nucleotide polymorphisms (SNPs) and insertions/deletions (InDels)) were identified using the three paired-end libraries for APS4 and the paired-end library for APS6, filtered as described above. Cleaned reads were aligned to the chromosomal assembly using bwa (version 0.7.12-r1039) (Li and Durbin 2009) and the resulting SAM alignments were converted to BAM format and sorted by coordinate using Picard (version 2.14) SortSam, deduplicated using picard MarkDuplicates and the BAM files indexed using picard BuildBamIndex. The three APS4 libraries were merged prior to deduplication and indexing. Joint variant calling was performed using Samtools mpileup (version 1.4) (Li et al. 2009) and the raw BCF output was filtered using bcftools view (version 1.4-16-g4dc4cd8) (Li 2011) and vcftools vcf-annotate (version 0.1.14, “-f +/d= 5/D= 10000/q= 20/Q= 15/w= 20/W= 30/c= 3,10/a= 2/1= 0.0001/2= 0/3= 0/4= 0.0001”) (Danecek et al. 2011). Intra-strain variants were defined as heterozygous polymorphisms occurring within one strain regardless of polymorphism at the same locus in the other strain. Inter-strain polymorphisms were defined as different genotypes

between the two strains at the same locus. Intra- and inter-strain variant density was plotted along each chromosome using KaryoploteR (Gel and Serra 2017). The gene and variant densities of unanchored scaffolds (i.e. those not mapped to a linkage group) were not examined.

Gene ontology (GO) enrichment analyses were performed to examine possible GO terms found over- or under-represented in the X chromosome gene set versus the autosomal one. For each enrichment analysis, we used a two-tailed Fisher's exact test (FDR < 0.05) implemented in the program Blast2GO (version 4.1.9) (Götz et al. 2008). The list of GO terms found enriched or depleted in the test set was then reduced to the most specific terms.

RNA extractions and transcriptomic analyses

L2 females, converted females and hermaphrodites. Female- and hermaphrodite-fated L2 larvae were isolated by using synchronized progeny populations generated by hermaphrodite mothers. Briefly, dauers (fated to become hermaphrodites) were isolated and allowed to develop into adult hermaphrodites. After ~12 h of egg laying the mothers were removed and the early eggs laid were left to hatch and grow until the L2 stage. During the L2 stage, females and hermaphrodites are distinguishable by their size and coloring: hermaphrodites are smaller, develop slower, are thinner and darker than females. Additionally, the female and hermaphrodite gonads are different in size during the mid-L1 stage, the female gonad being larger than that of the hermaphrodite. To convert hermaphrodite fated larvae into females, mid-L1 larvae with smaller gonads were isolated and grown with OP50-1 containing 200 nM dafachronic acid (DA) on NGM in individual wells of a 12 well culture plate. Once the larvae reached the L2 stage, the

ones that had a female morphology were collected and used for RNA extraction. About 200 L2s of each sex were picked and transferred to an Eppendorf tube containing 200 µL of M9 buffer (Stiernagle 2006). The nematodes were washed 2-3 times in M9 buffer, allowing them to sink to the bottom of the tube under gravity between each wash. After the final wash, the maximum amount of M9 was removed and 200 µL of Trizol was added to the tube. The tube was placed at -80 °C immediately. For RNA extraction, nematodes were first freeze-cracked in liquid nitrogen (2-3 times). Trizol was added to make up the volume to 500 µL and nematodes were shaken with a few sterile 0.5 mm glass beads on a BeadBeater homogenizer (20 s, 3 times with 30 s intervals). Subsequently, RNA was extracted using a standard chloroform approach and the pellet was dissolved in DEPC treated water and stored at -80 °C until further use.

Males and mixed stages samples. The same protocol was used to extract RNA from adult males and animals from various stages (mixed stages). To obtain RNA from males, about 500 young adult males were picked in a 1.5 mL centrifuge tube containing 0.5 mL of M9 and washed twice with the buffer. M9 was then replaced with 0.5 mL Trizol and the tube frozen at -80 °C. For the mixed staged nematodes, M9 was gently added to 5 culture plates (6 cm) containing a healthy population of nematodes. We avoided disturbing the bacterial lawn and naturally let the nematodes start to swim in the buffer. This method reduced bacterial contamination in the samples during harvesting. These nematodes were then collected into a centrifuge tube washed twice with M9, and then frozen using liquid nitrogen. Tissues were homogenized for 1 min using a probe homogenizer. After chloroform extraction, the RNA was dissolved in DEPC treated water and stored at -80 °C.

We generated three biological replicates for each RNA-seq condition (L2 females, L2 converted females, L2 hermaphrodites, males and mixed stages). RNA-seq was performed on the Illumina HiSeq2500 platform, generating a mean of 19.7 million 100 base read pairs per replicate. General assessment of the RNA-seq libraries was performed using FASTQC (Andrews 2010). The raw reads from each library were preprocessed using Trimmomatic (version 0.36, “HEADCROP:15 SLIDINGWINDOW:5:20 MINLEN:20”) (Bolger et al. 2014). The RNA-seq aligner STAR (version 2.4.2, “--sjdbOverhang 84”) (Dobin et al. 2013) was used to align the processed reads of each library to the primary scaffolded genome assembly. Transcript abundances were obtained using FeatureCounts (from the SubRead Package, version 1.5.0-p2) (Liao et al. 2014). Differential expression between the L2 females, L2 converted females and L2 hermaphrodites (three comparisons) was assessed using the R package DEseq2 (version 1.18.1, (Love et al. 2014)) following the standard procedure and generating diagnostic plots (as described in the DESeq2 documentation at https://bioconductor.org/packages/release/bioc/vignettes/DESeq2/inst/doc/DESeq2.html). An adjusted P-value of 0.01 and an absolute log2 fold change (FC) of 2 were used to define differentially expressed (DE) genes. Fold change of DE genes were plotted along the chromosomes using the package karyoploteR after lifting over annotations. Gene ontology (GO) enrichment analyses were conducted on the “down-regulated” and “up-regulated” genes separately for each comparison using the procedure explained above. Homologues of known sex determination genes were identified through reciprocal best hit BLASTp searches (BLAST+, “-evalue 0.01 -max_target_seqs 100, -outfmt 6”) using the *A. rhodensis* and *C. elegans* proteomes and analysed manually.

Global protein-coding gene expression of L2 females, L2 hermaphrodites, adult males and mixed stages was examined separately for each chromosome. The transcript abundances obtained using FeatureCounts (from the SubRead Package, version 1.5.0-p2) (Liao et al. 2014) were corrected by library size. The log2 of the global gene expression of each chromosome was plotted using ggplot2 in R. To determine if the genes of the X chromosome were significantly less expressed than those on the autosomes, we randomly sampled the same number of autosomal genes and X genes (600) and compared the sets using a Kruskal-Wallis test followed by Wilcoxon-Mann-Whitney tests between autosome-X pairs. The gene expression across the same chromosome in different replicates of the same condition was confirmed to be similar by performing Kruskal-Wallis tests.

RNAi of the DM domain gene Arh.g5747

The DM domain gene Arh.g5747 was significantly up-regulated in hermaphrodites compared to females in both female and converted female samples at L2. To further investigate this sex-specific expression of Arh.g5747 we targeted the gene for down-regulation using RNA interference (RNAi). Target specific dsRNA was produced using a cDNA template. PCR amplification was performed using the following primers (Forward primer: 5'-*TAATACGACTCACTATAGGG*TCATCAACGAGCAGAGCCGAGA-3', reverse primer: 5'-*TAATACGACTCACTATAGGG*TCCGCCTTCAGTGTTGGAGCT-3') to amplify an 858 bp fragment of the transcript of Arh.g5747. The T7 promoter (shown in italics above) was included at the 5' end of each primer to allow *in vitro* dsRNA synthesis. RNA extraction and cDNA synthesis were performed on ~300 adult hermaphrodite individuals as detailed above, with the exception that samples were subjected to

repeated cycles of freeze-thawing instead of bead-beating. RNA was treated with DNase I (Sigma) to remove residual genomic DNA. cDNA synthesis was performed with 0.5 µg of RNA using random primers (Promega) and the MMLV reverse transcriptase enzyme (Promega) following the manufacturer's instructions. The Arh.g5747 cDNA was then PCR amplified using GoTaq Green Master Mix (Promega), using approximately 200 ng of cDNA and 250 nM of each primer. PCR conditions were: an initial denaturation at 94 °C of 7 min, followed by 30 cycles of 94 °C for 15 s, 55 °C for 30 s and 72 °C for 60 s and a final extension of 10 min at 72 °C. After verification of the product size by gel electrophoresis, the amplicon was cleaned using the QIAquick PCR Purification kit (Qiagen), according to the manufacturer's protocol and eluted in a final volume of 25 µL. dsRNA was *in vitro* transcribed by incubating approximately 200 ng of the cDNA template with 2 µl (40 U) of T7 polymerase (Promega), 20 µl of 5x T7 polymerase buffer (Promega), 10 µl of DDT (Promega), 2.5 µl (100 U) of rRNAsin (Promega), 20 µl of 2.5 mM rNTPs (Thermo Fisher Scientific) and RNase free water to a final volume of 50 µl, for 4 h at 37 °C. The dsRNA product was size verified by gel electrophoresis and cleaned using the RNA clean-up protocol in the RNeasy® Mini Kit (Qiagen). A mixture of short interfering RNAs (siRNA) was produced by digesting 5 µg of the Arh.g5747 dsRNA with ShortCut RNase III (NEB) for 20 min at 37 °C, cleaned by glycogen/ethanol precipitation and eluted in 20 µl of RNase free water, according to the manufacturer's instructions. The RNAi mixture was produced by combining 5 µl of Arh.g5747 siRNA (approximately 100 ng), 4 ul of M9 buffer (Stiernagle 2006) and 1 µl (10% v/v) Lipofectamine®RNAiMax reagent (Invitrogen) and incubating at 25 °C for 20 min. Previously, we have shown that inclusion of the transfection reagent Lipofectamine dramatically improves RNAi efficiency in *A. rhodensis* (Adams et al. 2019). For control

injections Arh.g5747 siRNA was omitted from the mixture and replaced with additional M9 buffer.

For RNAi, young hermaphrodites (day 1 of adulthood) were immobilized on dried 2% agarose (w/v) pads in a small drop of undiluted halocarbon oil 700 (Sigma). The required injection mixture was loaded into pre-pulled microcapillary needles (Tritech Research) and microinjected into the gonad arms using an IM-300 Pneumatic Microinjector system with an oil hydraulic Micromanipulator (Narishige) using an injection pressure of 20 psi. Injected worms were rescued by adding a drop of M9 to the slide and moving them separately to a fresh 6 cm NGM plate seeded with *E. coli* OP50-1. The self-progeny from the injected hermaphrodites was sexed throughout the life of the mother. The mother was placed on a new plate every 24 h. Sex was determined according to the developmental rate, coloration and morphology of the larvae. Females were larger and whiter than hermaphrodites (dark and thin) due to their faster development. Males displayed a characteristic blunt tail.

Data Availability: Strains are available upon request. Sequence data are available at ENA and accession numbers are listed in Table 1. The genome assembly has been submitted to ENA under the accession number ERS3049325 (SAMEA5241922). Supplemental materials are available at Figshare.

Results

Genome characteristics

The scaffolded genome assembly spans 60.6 Mb in 440 scaffolds longer than 1000 bp. This span is smaller than that of *C. elegans* (100.2 Mb) but similar to that of other rhabditine nematodes (*Heterorhabditis bacteriophora*, 76.8 Mb; *Oscheius tipulae*, 59.0 Mb; *Caenorhabditis sulstoni*, 65.1 Mb). We predicted 11,570 protein coding genes and 833 unique non-coding RNAs. Previously sequenced rhabditine nematodes have been predicted to have more protein coding genes (*C. elegans,* 20,082; *H. bacteriophora*, 15,701; *O. tipulae*, 14,650). We presume that the difference in coding gene content is not due to a large number of missed genes, as the current assembly contains 99.19% of the core eukaryotic genes (predicted by CEGMA). The top BLAST hits of *A. rhodensis* proteins were more likely to be from parasitic Strongylomorpha species (*Ancylostoma ceylanicum*, *Haemonchus contortus* and *Necator americanus*) than from *C. elegans*. This is consistent with molecular phylogenies derived from ribosomal RNA and small numbers of protein coding loci (Kanzaki et al. 2017). However, it conflicts with analyses based on larger protein-coding gene datasets, which group *A. rhodensis* and the free-living *O. tipulae* are closer to *Caenorhabditis* species than to Strongylid species. A majority of the proteins (10,449, 90%) was assigned at least one type of annotation (InterProScan signature, GO term, BLAST hit) and 8181 (70.7%) were assigned with at least one GO term. The non-coding RNAs included all the expected major classes (Table S1). We also assembled and annotated the circular, 13,907 base pair mitochondrial genome (Figure S1). The genome assembly has been submitted to ENA under the accession number ERS3049325 (SAMEA5241922).

Genetic map

We generated 95 F2-derived progeny Lines (F2Ls) from crosses between two polymorphic inbred strains of *A. rhodensis* (strains APS4 and APS6). We identified 1,052 polymorphic RAD-seq markers that clustered in 7 linkage groups (Table 3), presumably corresponding to the seven chromosomes in *A. rhodensis* identified by DAPI staining (Tandonnet et al. 2018). We anchored the genomic scaffolds of *A. rhodensis* to the genetic map to complete the sequence of each linkage group. Of the 1,052 markers and 636 scaffolds (> 200 bp), 1,038 markers (~94%) and 143 scaffolds (~22%) were used to build the chromosomal assembly. The excluded scaffolds were generally short, and either lacked a RAD marker or had a marker that was not able to be placed. The anchored scaffolds represent 95.3% of the span of the scaffold span and contain 93.8% of the predicted proteins (Table 3).

Macrosynteny with *C. elegans* and identification of the X chromosome

A. rhodensis chromosomes are smaller than those of *C. elegans* (mean 8.3 Mb for *A. rhodensis* compared to 20.1 Mb for *C. elegans*), and one chromosome, LG5, is less than half the average size (3.5 Mb). To explore the origins of the changed complement of chromosomes and reduced size, we aligned the *A. rhodensis* protein-coding gene set to the one of *C. elegans* (Figure. 2). While there was a minor background of between-chromosome translocation, most chromosomes had congruent gene sets. The majority of loci on three *A. rhodensis* linkage groups (LG1, LG6 and LG7) mapped to single *C. elegans* chromosomes (V, IV and II, respectively), suggesting one-to-one chromosomal correspondence. For *A. rhodensis* LG2, LG3 and LG4 we observed mapping to two or

more chromosomes. Thus LG2 of *A. rhodensis* combines segments of *C. elegans* chromosomes III and X, LG3 combines segments of I and III, and LG4 combines segments of I, III and X. LG5, the smallest *A. rhodensis* chromosome, mapped almost entirely to the X chromosome of *C. elegans*, but segments of the *C. elegans* X chromosome were also found on LG2 and LG4 (Figure 3).

We confirmed that *A. rhodensis* LG5 was the X chromosome by genotyping F1 hybrid APS4/APS6 males at polymorphic loci spread across the linkage groups (Tandonnet et al. 2018). Hybrid males were heterozygous for all inter-strain polymorphic loci with the exception of those loci located in LG5, which were hemizygous in males (Tandonnet et al. 2018).

Chromosomal rearrangements

To further understand chromosomal rearrangements that took place in the lineage leading to *A. rhodensis*, we analysed the synteny relationships of loci conserved between *A. rhodensis* linkage groups and the chromosomal assemblies of *C. elegans*, *H. contortus, O. tipulae* and *Pristionchus pacificus* (Figures 1 and 3, Figure S2).

A. rhodensis LG1 (Ar_LG1) contained loci that had orthologues on a single chromosome in *C. elegans* and *H. contortus* (chromosomes V/5; Ce_V, Hc_5) and on the left arm of *P. pacificus* chromosome I (Pp_IL) (Figure 2 and Figure S2). Comparing to *O. tipulae*, the orthologues of loci on Ar_LG1 were on chromosome X (Oti_X) (Figure S2). *A. rhodensis* LG6 and LG7 had similar single-chromosome counterparts in the other species analysed. Orthologues of loci on Ar_LG6

were found on Ce_IV, Ot_IV, Hc_4 and Pp_IV. Orthologues of loci on Ar_LG7 were found on Ce_II, Ot_II, Hc_2 and Pp_I (Figure 2 and Figure S2).

Orthologues of loci on Ar_LG2 were found on two distinct chromosomes in *C. elegans*, *O. tipulae* and *H. contortus* (Ce_III, Hc_3, Ot_III and Ce_X, Hc_X, Ot_V). On Ar_LG2, these loci were partially segregated into blocks with different chromosomal locations in the other species (Figure 4, first column). In *P. pacificus*, these same blocks of loci had orthologues segregated on the right arm of chromosome 1 (Pp_IR) and on Pp_III. We concluded that Ar_LG2 was the product of fusion and rearrangement of a fragment or fragments of an ancestral chromosome represented by Pp_IR, Ot_V, part of Ce_X and part of Hc_X and an ancestral chromosome now present as parts of Pp_III, Ot_III, Ce_III and Hc_3. The presence of interspersed, extended blocks of loci that appeared to derive from the same ancestral chromosome suggested that the rearrangement was relatively recent, as the processes of intrachromosomal inversion, known to be very rapid in rhabditine nematodes (Stein et al. 2003), had not yet mixed up these blocks of genes.

Analyses of Ar_LG3 and Ar_LG4 identified similar patterns of breakage and fusion. Ar_LG3 contained two sets of distinct blocks of loci, one that had orthologues on Ce_III, Ot_III, Hc_3, and Pp_III, and a second that had orthologues on Ce_I, Ot_I, Hc_1, and Pp_V (Figure 4, middle column). Ar_LG4 had three sets of blocks of loci with orthologues on three chromosomes in other species: one set on Ce_III, Hc_3, Ot_III and Pp_III, one on Ce_X, Hc_X, Ot_V and Pp_ IR, and one on Ce_I, Hc_I, Ot_I and Pp_V (Figure 4, last column).

The evolutionary history of the X chromosome

The *A. rhodensis* X chromosome (LG5X) was the smallest chromosome (3.6 Mb) and had the lowest number of protein-coding genes (604, 5.5% of the total). X chromosomes differed markedly between species, but each contained orthologues of loci found on Ar_LG5X. For example, when compared to the most distantly related species, *P. pacificus* (Figure 5), the majority (82%) of the homologues of genes on Ar_LG5X were found on Pp_X (Figure 5A, Figure S2C). However, Pp_X was five times as big (16 Mb) and contained 2,998 protein coding genes (11.7% of all predicted *P. pacificus* genes).

C. elegans and *H. contortus* X chromosomes shared a striking pattern of macrosynteny with *A. rhodensis*. Mapping homologues from the *C. elegans* X to *A. rhodensis*, there were three distinct blocks of synteny on Ar_LG4 (Figure 5D), and 5 blocks of synteny on Ar_LG2 (Figure 5D). These same blocks were observed in comparisons with *H. contortus* (Figure 5C). Intriguingly, while the *C. elegans* orthologues of Ar_LG5X, Ar_LG4 and Ar_LG3 loci were evenly distributed across Ce_X (Figure 5E), the mapping of the *A. rhodensis* orthologues on the *H. contortus* X was partitioned. Ar_LG4 matches were clustered on the left end of Hc_X and Ar_LG5X matches on the right (Figure 5E). Ar_LG2 matches were more evenly distributed, although we can note a clustering on both ends of Hc_X. As discussed above, Ar_LG2 and Ar_LG4 may have originated through chromosome breakage and rearrangement. The segregation of Ar_LG5X-like and Ar_LG4-like regions on the *H. contortus* X may reflect conservation of ancestral synteny that has not been homogenized by within-chromosome rearrangement. The contrast between Ce_X and Hc_X, two chromosomes that otherwise appear highly

homologous, suggested that either intrachromosomal rearrangement has been much more active in the lineage leading to *C. elegans* or that *A. rhodensis* and *H. contortus* shared a more recent common ancestor. *A. rhodensis* orthologues of genes on the *O. tipulae* X chromosome were found on Ar_LG5X and Ar_LG1 (Figure 5B). This pattern was different from the one shared by mappings to *P. pacificus*, *H. contortus* and *C. elegans* and may reflect a novel trajectory of X chromosome evolution in the branch leading to *O. tipulae*.

Contrasting patterns of genome structure between the X chromosome and the autosomes

We explored large-scale patterns of genome structure and evolution across the *A. rhodensis* genome. In *C. elegans*, conserved genes are more frequently found in the centres of chromosomes and are rarer in autosomal chromosome arms (Wilson 1999; Consortium 1998). However, in *A. rhodensis* the gene density and localisation of genes with orthologues in *Drosophila melanogaster* across chromosomes was uniform (Figure 6A). LG5X had a lower gene density than the autosomes (one protein-coding gene per 5.8 kb on LG5X compared to 5.3 kb ± 0.2 kb on the autosomes), and fewer conserved genes were present on LG5X (Figure 6A, Table S2). Most strikingly, none of the nearly 500 tRNA loci were on LG5X (Figure 6B).

A gene ontology analysis comparing the X *versus* the autosomal gene sets found that the GO terms ‘translation’, ‘ribosome’, ‘nucleic acid binding’, ‘intracellular membrane-bounded organelle’ and ‘hydrolase activity’ were under-represented on LG5X compared to the autosomes (Table S3). The process governing the ‘neuropeptide signaling pathway’ was found enrich on LG5X (Table S3).

Although the strains used in this study were inbred, we expected to observe a low level of within-strain heterozygosity due to incomplete inbreeding. We found that strain APS6 had higher heterozygosity than APS4, probably because APS6 underwent less inbreeding than APS4 (11 rounds of bottlenecking *versus* 50) (Figure 6C, Table S4). While the overall frequency of variants was different in the two strains, the distribution of these variants across the genome was similar (Figure 6C), including shared chromosomal regions with higher natural variability. The LG5X chromosome displayed more within-strain variation than the autosomes, probably due to the atypical inheritance of this chromosome.

The genetic map displayed deviation from expected Hardy-Weinberg equilibria in several regions. We found that almost all RAD markers for LG5X were heterozygous across all 95 samples (Figure 6D). This distorted pattern of X heterozygosity can be explained by the fact that the RAD data were derived from F2 hermaphrodite progenitors left to propagate for 3-10 generations, and there is no X recombination in hermaphrodites (Tandonnet et al. 2018). The few (~10%) markers that were homozygous for the X chromosome are probably the result of recombination that occurred in females during the population expansions originating from the F2s. We also observed a high frequency of homozygous markers for APS6 alleles at the right end of LG3 (Figure 6D), suggesting that APS6 alleles had been positively selected in the culture conditions tested or possibly that segregation distorters were present. Less extreme deviation from expected equilibrium was also observed at the left end of LG7. These deviations were not explored further.

The different dosage of X chromosomes in females and hermaphrodites compared to males results in a requirement for dosage compensation for X-linked genes. We examined global protein-coding gene expression in L2 females, L2 hermaphrodites, adult males and mixed stages and compared autosomal and X-linked genes (Figure 7). After correction for library size, gene expression from each autosome was found to be similar both between autosomes and across lifecycle stages. However, and unlike *C. elegans*, genes on *A. rhodensis* LG5X showed consistently lower expression than those on the autosomes, even in the L2 stage (Wilcoxon Mann-Whitney, p-value <= 1.0e-11 in all conditions and replicates).

Transcriptomic identification of loci associated with sexual morph development

We compared gene expression in developing XX L2 larvae to identify loci that may be associated with the different sexual morphs of *A. rhodensis*. We generated replicate RNA-seq datasets from L2 fated to become hermaphrodites, L2 fated to become females, and L2 from hermaphrodite-fated nematodes that were converted to females by treatment with DA (converted females). Using standard thresholds (absolute log2(Fold Change) >= 2, FDR <0.01), we found 2,422 (21%) of the predicted genes were differentially expressed (DE) between L2 hermaphrodites and L2 females. Slightly fewer genes (2,121,18%) were DE between L2 hermaphrodites and L2 converted females. Most of the genes found to be DE between females and hermaphrodites and between converted females and hermaphrodites were the same (Figure 8A). The genes more expressed in females and converted females compared to hermaphrodites were enriched in GO terms related to translation, protein synthesis, ribosomal function, gonad

609 and embryo development and structural constituents ofthe cuticle (see File S1
610 for the list of complete terms). Genes more expressed in hermaphrodites were enriched
611 in few GO terms, with only “structural constituent of cuticle” in common between both
612 comparisons. DE genes were distributed across the *A. rhodensis* genome, with no
613 enrichment or depletion on LG5X (Figure 8B).

614 Some *A. rhodensis* orthologues of *C. elegans* sex determination genes were DE
615 between female (normal or converted) and hermaphrodite L2s. The known sex
616 determination genes Arh.g5696-*gld-1* and Arh.g4999-*tra-1* were 200 and 4 times more
617 expressed in females, respectively. The precise roles of *tra-1* and *gld-1* in the sex
618 determination in *A. rhodensis* are not yet known. A number of *daf* (dauer formation)
619 genes (Arh.g6122-*daf-11*, Arh.g7695-*daf-16*, Arh.g7696-*daf*-like) were expressed at
620 higher levels in hermaphrodite L2 compared to female or converted female L2,
621 consistent with the obligate transition through dauer of hermaphrodite-fated nematodes.

622 The DM (*doublesex*/*mab-3*) domain transcription factor Arh.g5747 (*dmd-10*/*11*-like) was
623 found to be more than 200 times more expressed in hermaphrodite L2 than in female or
624 converted female L2. To investigate the role of this locus in the decision between
625 female and hermaphrodite sexual fate in *A. rhodensis*, we downregulated Arh.g5747 by
626 injecting RNAi in young hermaphrodites (first day of adulthood). If Arh.g5747 is required
627 for determining hermaphrodite fate, we would expect to see more female progeny from
628 injected hermaphrodite mothers. Indeed, downregulation of Arh.g5747 in 8
629 hermaphrodite mothers resulted in more female progeny than control injections
630 performed on 9 hermaphrodites (Wilcoxon Mann-Whitney test, W = 67, p-value =
631 0.001563, Table 4). Thus, this DM domain locus may drive *A. rhodensis* hermaphrodite

fate, either by inhibiting a female induction signal or through positive upregulation of a hermaphrodite-inducing pathway.

The comparison of the L2 females and the L2 converted females is particularly interesting for identifying genes or mechanisms involved in the hermaphrodite-female decision, upstream of the DA pathway. Gene expression in L2 females and L2 converted females was strikingly similar. Only 55 genes (0.5%) were found to be more highly expressed in females compared to converted females (Figure 8A and 7B). No genes were found to be significantly less expressed in females compared to converted females. This result is surprising, since the female-inducing treatment (DA) was applied at the L1 stage, when sexual fate has already been decided, and the transcriptome was sampled less than 24 h after DA application. Functional annotation of these DE genes revealed several whose *C. elegans* homologues are involved in embryogenesis and developmental processes. Three chondroitin proteoglycan genes (Arh.g2548, Arh.g5439, Arh.g2211) were more expressed in female L2, and were also DE between female L2 and hermaphrodite L2. Chondroitin proteoglycans are important for embryonic cell division and vulval morphogenesis. More strikingly, we identified the homologues of the zinc-finger genes *mex-1* (required for germ cell formation, and somatic cell differentiation in the early embryo in *C. elegans*) and *pos-1* (essential for proper fate specification of germ cells, intestine, pharynx, and hypodermis in *C. elegans*). Both these genes had very low expression in converted female L2 and hermaphrodite L2. As these genes are maternally supplied in *C. elegans*, one possibility is that they are maternal regulators of sexual fate in *A. rhodensis*, although this hypothesis remains speculative.

Discussion

Auanema rhodensis is a rare example of a three-sexed animal. Here we sequenced the *A. rhodensis* genome and used a linkage map to construct a chromosomal assembly. At 60 Mb, the *A. rhodensis* genome is smaller than that of *C. elegans* (100 Mb), but within the range (55-160 Mb) of other free-living rhabditomorph nematodes. We predict only 11,570 protein coding genes, many fewer than the 23,000 identified in *C. elegans*, and fewer than would be predicted from the reduction in genome size alone. However, considering that ~99% of the core eukaryotic genes (CEGMA prediction) were identified, the reduced gene count in *A. rhodensis* is unlikely due to a high number of unannotated genes.

It is known that mating system can influence genome size. In the *Caenorhabditis* clade, selfing (hermaphrodite) species have smaller genomes than their outcrossing sister species (Yin et al. 2018; Fierst et al. 2015), but these differences are in the order of 10%. However, other free-living and entomopathogenic rhabditomorphs have genomes smaller than *C. elegans*, and the related animal-parasitic Strongylomorpha have much larger genomes (250 – 700 Mb). Detailed understanding of the evolutionary drivers of genome size in this group awaits additional, dense sampling across the Rhabditomorpha.

While there is little shared gene order between nematode species, we identified strong macro-syntenic patterns between *A. rhodensis, P. pacificus, H. contortus, O. tipulae* and *C. elegans*. These patterns allow us to propose a preliminary model of the evolution of chromosomes in the Rhabditina, which includes Diplogasteromorpha (*P. pacificus*), Strongylomorpha (*H. contortus*), and Rhabditomorpha (*A. rhodensis, O. tipulae* and *C.*

678 *elegans*). It has long been noted that the majority of rhabditine nematodes have a
679 karyotype of n= 6, with an XX:XO sex chromosome system (Walton 1959). While there
680 are deviations from this pattern, including the trioecious *A. rhodensis*, with n= 7, and the
681 variously parthenogenetic *Diploscapter* species with n=1 to n=9, the phylogenetic
682 perdurance of this karyotype is striking. Using loci identified as orthologues in each
683 species pair, we could identify six putative ancestral macrosynteny groups (Figure 9)
684 and also map the macrosyntenic changes that may have given rise to present day
685 karyotypes.

686 We call these ancestral linkage groups Nigon elements in homage to Victor Nigon
687 (Nigon and Felix 2017), a name coined by Matt Rockman (personal communication) in
688 analogy with the Muller units of *Drosophila* chromosomes. Some of these Nigon
689 elements have been transmitted intact through Rhabditina, while others have
690 undergone rearrangement. Where a rearrangement has resulted in the fusion of
691 (parts of) Nigon elements, in most cases the dynamic processes of intrachromosomal
692 rearrangement, which are very active in rhabditine nematodes (Stein et al. 2003), have
693 acted to mix up the genes originally derived from different units. In other cases, either
694 because the fusions were more recent or because the processes of intra-chromosomal
695 rearrangement are less active, the sets of loci from distinct Nigon elements are found as
696 blocks in the fusion chromosome. We define six Nigon elements, NA, NB, NC, ND, NE
697 and NX, as well as an additional NN unit which we are currently unable to place (Figure
698 9). It could originally link to NE or NX, or be a unit of its own (which would imply seven
699 Nigon elements in total).

A. rhodensis LG1, LG6 and LG7 represent chromosomal units that have survived largely intact through rhabditine evolution (Nigon elements NE, ND and NB, respectively). However, the *A. rhodensis* X chromosome (LG5X) was formed from a subset of the loci now found on the *C. elegans* X chromosome, and *A. rhodensis* LG2, LG3 and LG4 are the products of major interchromosomal rearrangement events. For example, NA is intact in *C. elegans* (Ce_I), *H. contortus* (Hc_1), *O. tipulae* (Ot_I) and *P. pacificus* (Pp_V), but underwent fission in the *A. rhodensis* lineage, with subsequent fusion forming two hybrid chromosomes. *A. rhodensis* LG3 is formed largely from loci from part of NA (subset a1) and NC (c2), while LG4 includes loci from NA (a2), NC (c3) and NN (n2). Overall, compared to the other four species with chromosomal assemblies (or chromosome-allocated scaffolds), the fission/fusion event(s) may be directly associated with the origin of the novel n=7 karyotype of this species. It will be informative to explore the origins of other species where n≠ 6, such as species in the genus *Diploscapter*. In *A. rhodensis*, these events may have been relatively recent, phylogenetically speaking, as there are still clear blocks of genes of different Nigon element origin within the fusion chromosomes LG2, LG3 and LG4.

Fusions are not unique to *A. rhodensis*. NE is intact in *A. rhodensis* (Ar_LG1), *C. elegans* (Ce_V) and *H. contortus* (Hc_V) but has fused with NN (n1n2) in *P. pacificus* to form Pp_I. As noted previously (Rödelsperger et al. 2017), Pp_I is a fusion chromosome incorporating components of Ce_V (NE-derived) and Ce_X (NN-derived), but we note that the continued distinctiveness of the NE-derived and NN-derived components within Pp_I suggests that this fusion is phylogenetically recent. The two Nigon element components within Pp_I retain an arms-and-centres long-range structure that is presumably derived from the original separate chromosomes, with high repeat density

724 in the ancestral arms and high gene density in the ancestral centres. It has been
725 proposed that the NE-NN fusion observed in *P. pacificus* is ancient, based on
726 identification of a NE-NN like junction fragment in the genome sequence of the
727 tylenchine (Clade IV) nematode *Bursaphelenchus xylophilus*, which is an outgroup to
728 the rhabditine species. This apparent conservation of the junction fragment conflicts
729 with the within-chromosome rearrangements dynamic observed elsewhere in the
730 genome, and may be a chance homoplastic association of NE and NN elements in this
731 species. In *O. tipulae*, NE has fused with NX to form Ot_X, and the NN unit forms a
732 chromosome on its own (Ot_V).

733 The X chromosomes of the species analysed always contain the NX unit either as the
734 sole component of the X (Pp_X and Ar_LG5X), or associated with other Nigon
735 elements: NN in Ce_X and Hc_X, and NE in Ot_X. The complex history of the NN, NX
736 and NE units requires additional analyses, as it is unclear if NN belongs to NE (as found
737 in Pp_I), to NX (as found in Ce_X and Hc_X) or if it is a unit of its own (as found in
738 Ot_V). The number of chromosomally-assembled rhabditine genomes is still too few to
739 fully define the ancestral gene content of Nigon elements.

740 Meiosis of the *A. rhodensis* X chromosome seems to follow different patterns, largely
741 depending on the organismal sex and type of gametogenesis. By using five polymorphic
742 markers, we have previously shown that the homologous X chromosomes undergo
743 meiotic recombination in females, but not in hermaphrodites (Tandonnet et al. 2018;
744 Shen and Ellis 2018).The genetic linkage map performed in the present study, using 92
745 markers for the X chromosome, confirms that the lack of recombination in the X in
746 hermaphrodites is chromosome-wide (Figure 6D). Lack of recombination of the X is

observed during hermaphrodite oogenesis and spermatogenesis, leading to nullo-X oocytes and diplo-X sperm (Tandonnet et al. 2018). Additionally, during outcrossing the X chromosome is always transmitted from father to son (males produce exclusively haplo-X sperm). One of predictions from this atypical inheritance is that the genes on the X will be more exposed to selection. Thus, essential genes will tend to migrate from the X to autosomes, leading to a reduction in size. The X chromosome of *A. rhodensis* has many distinctive features. It is much smaller than the autosomes, representing only 6% of the genome and containing just over 600 genes. It has no tRNA genes, and fewer conserved genes were found on the X compared to the autosomes. In *C. elegans,* the X chromosome carries 44% of all tRNA genes and there is no marked exclusion of conserved genes (Wilson 1999). In *A. rhodensis*, the X is inherited from father to son and is haploid in males. Thus, genes on X chromosomes transmitted between males will be more exposed to natural selection, which will tend to exclude essential genes from the X (Tandonnet et al. 2018). In addition, the lack of recombination of the X chromosome in hermaphrodites will slow down the removal of deleterious mutations, contributing to the exclusion of essential genes on the X. The lower prevalence of essential genes on the X chromosome was reported in *C. elegans* (Kamath et al. 2003), where a genome-wide RNAi analysis revealed that the X chromosome was depleted of essential genes. The *C. elegans* X chromosome, which is also haploid in males, would also be more exposed to natural selection than autosomes, although to a lesser degree than *A. rhodensis*. The heightened exposure to selective forces and lack of recombination of the X would predict a lower diversity on the X due to genetic hitchhiking. Indeed, the presence of a beneficial allele on the X would quickly spread through the population drawing along the rest of the chromosome (selective sweep),

771 and, correspondingly, the negative selection of a deleterious allele on the X would also
772 lead to a decrease of the genetic diversity of the whole X (background selection).
773 However, populations of *A. rhodensis* are composed of a high proportion of selfing
774 hermaphrodites (estimated around 60% of the adult fraction of the population), in which
775 the X chromosome does not recombine. The XX progeny resulting from a selfer
776 therefore always retain maternal heterozygosity on the X (Tandonnet et al. 2018). Using
777 the inbred strains APS4 and APS6, we found that within-strain and between-strain
778 genetic diversities were higher on the *A. rhodensis* X chromosome than on the
779 autosomes. In our inbreeding protocol, bottlenecking was performed by isolating a
780 single selfing hermaphrodite every few generations. Thus, the X chromosome will only
781 have recombined in females during the population expansion from each isolated
782 hermaphrodite. As several generations occurred between each hermaphrodite isolation,
783 we expect that the X would become homozygous at a much slower rate than the
784 autosomes. In nature, the genetic diversity of the X chromosome compared to the
785 autosomes (whether higher or lower) could depend on the proportion of the different
786 sexes and on the effect of X-linked mutations.

787 The X chromosome gene expression was also found to be consistently lower than that
788 of autosomes. In *C. elegans* hermaphrodites, the X to autosomal gene expression ratio
789 (X:A ratio) varies through development from 0.92 in the L2 to 0.41 fold in adults (Xiong
790 et al. 2010). This change is likely to be associated with the exclusion of genes with
791 germline expression from the *C. elegans* X chromosome and the growth of the germline
792 in adults (Gama et al. 2002; Strome et al. 2014). Indeed, as the individual (XO or XX)
793 develops, the proportion of germ cells increases. Because the X chromosome is
794 repressed in germ cells, the X:A ratio steadily declines as the individual develops (Deng

et al. 2011). However, in *C. elegans* mutants lacking germline proliferation, the X:A ratio is close to 1, due to a dosage compensation mechanism equalizing the X chromosome and autosomal gene expression (Deng et al. 2011). In *A. rhodensis*, the X chromosome expression was found lower than autosomal expression even at the L2 stage. Based on these observations, it is possible that the X chromosome of *A. rhodensis* lacks a dosage compensation mechanism to equalize the X:A ratio (unlike *C. elegans*). However, it is to note that the X expression seems similar between XO males and XX animals, indicating that some dosage compensation mechanism is acting to prevent a higher expression in XX animals compared to XO males. Considering the few gene number on the X, one possibility is that a low X:A ratio is viable in *A. rhodensis*.

Another fundamental question in *A. rhodensis* biology is the mechanism that controls female *versus* hermaphrodite sex determination in XX animals. Females and hermaphrodites are karyotypically identical, and are thought to be genetically identical. This is because hermaphrodites of a strain inbred for 50 generations (APS4) still produces hermaphrodite and female progeny (Chaudhuri et al. 2015). Transcriptome comparisons between female, hermaphrodite and converted female early larvae (L2) show that the sex differentiation process modulates the expression of many genes, with ~20% of all genes found differentially expressed between females (normal and converted females) and the hermaphrodites. The considerable difference in transcriptomic profiles at L2 is reflected in their developmental trajectory, with the female-fated larvae undergoing faster development towards adulthood, whereas hermaphrodite-fated larvae arresting at the dauer stage. Additionally, physiological and metabolic differences between females and hermaphrodites are likely to be present. One established example is the production of male-attracting pheromones by females,

but not in hermaphrodites (Chaudhuri et al. 2015). Of particular interest, we noted that *A. rhodensis* orthologues of some genes previously shown to be active in *C. elegans* sex determination were differentially expressed in females *versus* hermaphrodites. *A. rhodensis gld-1*, for example, was 200-fold overexpressed in females. In *C. elegans* hermaphrodites, *gld-1* is necessary for normal oogenesis and promotes spermatogenesis in hermaphrodites (Jan et al. 1999; Jones and Schedl 1995). However, it has the opposite role in *C. briggsae* (Beadell et al. 2011; Beadell and Haag 2014), probably due to changes in target transcripts. In *C. elegans*, GLD-1, and its cofactor FOG-2, regulate the translation of *tra-2* and are necessary for hermaphrodite sperm fate (Hu et al. 2019). In *C. elegans*, wild-type and *fog-2* mutants have very similar transcriptomic profiles, which is consistent their role at the translational level (Hu et al. 2019). In *A. rhodensis* however, no *fog-2* orthologue was found and the L2 female and hermaphrodite transcriptomes are strikingly different, indicating that the regulation of sexual fate is different from *C. elegans*. *A. rhodensis tra-1*, a master sex determination gene, was 4-fold overexpressed in females. The zinc finger protein TRA-1 is the terminal regulator of the sex determination pathway in *C. elegans*, where it promotes female development in somatic tissues (Hodgkin 1987; Zarkower and Hodgkin 1992). Its role in sex determination is conserved in nematodes, making it an interesting target for functional studies (Pires-daSilva and Sommer 2004). We also identified a DM (*doublesex*/*mab-3*) domain transcription factor, homologous to *C. elegans dmd-10* and *dmd-11* that was 200-fold overexpressed in hermaphrodites. DM domain genes regulate sex determination and sexual differentiation processes in a number of organisms (Zarkower 2013), but specific roles for *C. elegans dmd-10* and *dmd-11* have not yet been elucidated. RNAi knockdown of this locus in hermaphrodites resulted in the

production of more female progeny, suggesting a role for this DM domain-coding gene in *A. rhodensis* sex determination.

The near identical expression pattern between converted females and females shows that the conversion of L1 hermaphrodite-fated larvae by exposure to DA is almost complete and that the initial sex decision can be overridden almost fully by hormonal manipulation. The age and sex of an *A. rhodensis* mother affect the proportion of each sex in its progeny (Chaudhuri et al. 2015) and thus it is likely that maternal effects may directly establish the distinct developmental trajectories of females and hermaphrodites. However, the female *versus* hermaphrodite decision could also be modulated by environmental cues acting during embryogenesis and the L1 stage. These maternal and environmental effects could be modulations of what is essentially a random sex determination (RSD) system (Perrin 2016). RSD occurs when fluctuations in the expression of genes at the top of the sex determining cascade or “developmental noise” are enough to canalize sexual fate down contrasting paths. An RSD component in *A. rhodensis* is plausible as females and hermaphrodites likely share the same genome, all sexual morphs are produced in a single environmental condition and the proportion of each sex produced varies greatly between mothers (although they are from inbred lines).

Auanema nematodes thus offer a fascinating and potentially highly informative model system for depth exploration of the origin of novel traits and their consequences. The genomic and transcriptomic resources we present for *A. rhodensis* will be critical for future analyses of the origins of new chromosomes in an otherwise stable karyotypic system, the biology of the highly regulated pattern of X chromosome segregation, the

dynamics of mating system evolution, and the evolution of sex determination mechanisms. Towards this, we have identified the orthologs of 16 main sex determination genes, including *tra-1/2*, *gld-1*, *her-1* and *fem-1/2* (Table S5), which are clear candidates to investigate the sex determination mechanisms in *A. rhodensis*. In parallel, we are developing reverse genetic and functional genomic technologies for these species, and these promise routes to rapid validation of hypotheses of gene function (Adams et al. 2019). The *A. rhodensis* sex determination system may integrate genetic, maternal, environmental and random components, and this nexus of interacting components will also become amenable to manipulation and dissection. Genetic and genomic investigation of additional *Auanema* and closely related rhabditomorph species will contribute to a complete understanding of the origins and maintenance of this unusual mating system.

Author Contributions

Conceptualization: ST, APS and MB; Generation of F2Ls, DNA and RNA extractions: MP; Genome assembly and annotation: GDK and ST; Genetic linkage map: ST; Genome analysis: ST, MB, APS; Transcriptomic analysis: ST; RNAi: SA, DC and ST; Data curation: ST and GDK; Supervising: APS and MB; Manuscript writing: ST, MB, APS and SA with input from all authors.

Acknowledgments

G.D.K. was supported by a BBSRC Ph.D. studentship and S.T. by a Ph.D. training grant from CAPES/CNPq (201116/2014-6). A.P.S. was supported by a grant from National Science Foundation (IOS 1122095), BBSRC (BB/L019884/1) and University of Warwick start-up funds. The authors are very grateful to Stephen Doyle (from the Sanger Institute) for providing early access to the *Haemonchus contortus* genome and annotation, and to Fabrice Besnard for discussion about *O. tipulae* linkage groups. We thank staff of Edinburgh Genomics and UT Southwestern facilities for support in sequencing. We are also grateful to two anonymous reviewers for their constructive comments on an earlier version of the manuscript.

Literature cited:

Adams, S., P. Pathak, H. Shao, J.B. Lok, and A. Pires-daSilva, 2019 Liposome-based transfection enhances RNAi and CRISPR-mediated mutagenesis in non-model nematode systems. *Sci Rep* 9 (1):483.

Amores, A., J. Catchen, I. Nanda, W. Warren, R. Walter *et al.*, 2014 A RAD-tag genetic map for the platyfish (*Xiphophorus maculatus*) reveals mechanisms of karyotype evolution among teleost fish. *Genetics* 197 (2):625-641.

Andrews, S., 2010 FastQC: a quality control tool for high throughput sequence data in www.bioinformatics.babraham.ac.uk/projects/fastqc.

Baird, N.A., P.D. Etter, T.S. Atwood, M.C. Currey, A.L. Shiver *et al.*, 2008 Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3 (10):e3376.

Baldi, C., S. Cho, and R.E. Ellis, 2009 Mutations in two independent pathways are sufficient to create hermaphroditic nematodes. *Science* 326 (5955):1002-1005.

Beadell, A.V., and E.S. Haag, 2014 Evolutionary Dynamics of GLD-1-mRNA complexes in Caenorhabditis nematodes. *Genome Biol Evol* 7 (1):314-335.

Beadell, A.V., Q. Liu, D.M. Johnson, and E.S. Haag, 2011 Independent recruitments of a translational regulator in the evolution of self-fertile nematodes. *Proc Natl Acad Sci U S A* 108 (49):19672-19677.

Bell, G., 1982 *The masterpiece of nature: the evolution and genetics of sexuality*. Berkeley: University of California Press.

Besnard, F., G. Koutsovoulos, S. Dieudonne, M. Blaxter, and M.A. Felix, 2017 Toward Universal Forward Genetics: Using a Draft Genome Sequence of the Nematode Oscheius tipulae To Identify Mutations Affecting Vulva Development. *Genetics* 206 (4):1747-1761.

Blaxter, M., and G. Koutsovoulos, 2014 The evolution of parasitism in Nematoda. *Parasitology*:1-14.

Bolger, A.M., M. Lohse, and B. Usadel, 2014 Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30 (15):2114-2120.

Brenner, S., 1974 The genetics of *Caenorhabditis elegans*. *Genetics* 77 (1):71-94.

Broman, K.W., H. Wu, S. Sen, and G.A. Churchill, 2003 R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19 (7):889-890.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos *et al.*, 2009 BLAST+: architecture and applications. *BMC Bioinformatics* 10:421.

Campbell, M.S., C. Holt, B. Moore, and M. Yandell, 2014 Genome Annotation and Curation Using MAKER and MAKER-P. *Curr Protoc Bioinformatics* 48:4 11 11-39.

Catchen, J., P.A. Hohenlohe, S. Bassham, A. Amores, and W.A. Cresko, 2013 Stacks: an analysis tool set for population genomics. *Mol Ecol* 22 (11):3124-3140.

Charlesworth, D., 1984 Androdioecy and the evolution of dioecy. *Biol J Linn Soc* 22 (4):333-348.

Charlesworth, D., 2006 Evolution of plant breeding systems. *Curr Biol* 16 (17):R726-735.

Charlesworth, D., M.T. Morgan, and B. Charlesworth, 1990 Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. *Evolution* 44 (6):1469-1489.

Charnov, E.L., 1982 *The theory of sex allocation*. Princeton, N.J.: Princeton University Press.

Chasnov, J.R., and K.L. Chow, 2002 Why are there males in the hermaphroditic species *Caenorhabditis elegans*? *Genetics* 160 (3):983-994.

Chaudhuri, J., N. Bose, S. Tandonnet, S. Adams, G. Zuco *et al.*, 2015 Mating dynamics in a nematode with three sexes and its evolutionary implications. *Sci Rep* 5:17676.

Chaudhuri, J., V. Kache, and A. Pires-daSilva, 2011 Regulation of sexual plasticity in a nematode that produces males, females, and hermaphrodites. *Curr Biol* 21 (18):1548-1551.

Chikhi, R., and P. Medvedev, 2014 Informed and automated k-mer size selection for genome assembly. *Bioinformatics* 30 (1):31-37.

Chitwood, B.G., and M.B.H. Chitwood, 1950 *An introduction to nematology*. Washington.

Cho, S., S.W. Jin, A. Cohen, and R.E. Ellis, 2004 A phylogeny of *Caenorhabditis* reveals frequent loss of introns during nematode evolution. *Genome Res* 14 (7):1207-1220.

Ciche, T., 2007 The biology and genome of *Heterorhabditis bacteriophora*. *WormBook*:1-9.

Consortium, T.*C.e.*S., 1998 Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282 (5396):2012-2018.

Cutter, A.D., 2005 Mutation and the experimental evolution of outcrossing in *Caenorhabditis elegans*. *J Evol Biol* 18 (1):27-34.

Danecek, P., A. Auton, G. Abecasis, C.A. Albers, E. Banks *et al.*, 2011 The variant call format and VCFtools. *Bioinformatics* 27 (15):2156-2158.

Darwin, C., 1876 *The effects of cross and self fertilisation in the vegetable kingdom*. London, UK: John Murray.

Deng, X., J.B. Hiatt, D.K. Nguyen, S. Ercan, D. Sturgill *et al.*, 2011 Evidence for compensatory upregulation of expressed X-linked genes in mammals, Caenorhabditis elegans and Drosophila melanogaster. *Nat Genet* 43 (12):1179-1185.

Dobin, A., C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski *et al.*, 2013 STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29 (1):15-21.

Félix, M.A., 2004 Alternative morphs and plasticity of vulval development in a rhabditid nematode species. *Dev Genes Evol* 214 (2):55-63.

Fierst, J.L., J.H. Willis, C.G. Thomas, W. Wang, R.M. Reynolds *et al.*, 2015 Reproductive Mode and the Evolution of Genome Size and Structure in Caenorhabditis Nematodes. *PLoS Genet* 11 (6):e1005323.

Gama, S.M., R. De Gasperi, P.H. Wen, E.A. Gonzalez, K. Kelley *et al.*, 2002 BAC and PAC DNA for the generation of transgenic animals. *Biotechniques* 33 (1):51-53.

Gel, B., and E. Serra, 2017 karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* 33 (19):3088-3090.

Glemin, S., 2007 Mating systems and the efficacy of selection at the molecular level. *Genetics* 177 (2):905-916.

Goodwillie, C., S. Kalisz, and C.G. Eckert, 2005 The evolutionary enigma of mixed mating systems in plants: Occurrence, theoretical explanations, and empirical evidence. *Ann Rev Ecol Evol Syst.* 36:47-79.

Götz, S., J.M. Garcia-Gomez, J. Terol, T.D. Williams, S.H. Nagaraj *et al.*, 2008 High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 36 (10):3420-3435.

Grabherr, M.G., B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson *et al.*, 2011 Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29 (7):644-652.

Herrmann, M., W.E. Mayer, and R.J. Sommer, 2006a Nematodes of the genus *Pristionchus* are closely associated with scarab beetles and the Colorado potato beetle in Western Europe. *Zoology (Jena)* 109 (2):96-108.

Herrmann, M., W.E. Mayer, and R.J. Sommer, 2006b Sex, bugs and Haldanes rule: the nematode genus *Pristionchus* in the United States. *Front Zool* 3 (1):14.

Hodgkin, J., 1987 A genetic analysis of the sex-determining gene, *tra-1*, in the nematode *Caenorhabditis elegans*. *Genes Dev* 1 (7):731-745.

Hodgkin, J., 2002 Exploring the envelope. Systematic alteration in the sex-determination system of the nematode *Caenorhabditis elegans*. *Genetics* 162 (2):767-780.

Hu, S., L.E. Skelly, E. Kaymak, L. Freeberg, T.W. Lo *et al.*, 2019 Multi-modal regulation of C. elegans hermaphrodite spermatogenesis by the GLD-1-FOG-2 complex. *Dev Biol* (2): 446, 193-205

Hunt, M., T. Kikuchi, M. Sanders, C. Newbold, M. Berriman *et al.*, 2013 REAPR: a universal tool for genome assembly evaluation. *Genome Biol* 14 (5):R47.

Jan, E., C.K. Motzny, L.E. Graves, and E.B. Goodwin, 1999 The STAR protein, GLD-1, is a translational regulator of sexual identity in *Caenorhabditis elegans*. *Embo J* 18 (1):258-269.

Jiang, H., R. Lei, S.W. Ding, and S. Zhu, 2014 Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* 15:182.

Jones, A.R., and T. Schedl, 1995 Mutations in gld-1, a female germ cell-specific tumor suppressor gene in *Caenorhabditis elegans*, affect a conserved domain also found in Src- associated protein Sam68. *Genes Dev* 9 (12):1491-1504.

Jones, P., D. Binns, H.Y. Chang, M. Fraser, W. Li *et al.*, 2014 InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30 (9):1236-1240.

Kamath, R.S., A.G. Fraser, Y. Dong, G. Poulin, R. Durbin *et al.*, 2003 Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* 421 (6920):231-237.

Kanzaki, N., K. Kiontke, R. Tanaka, Y. Hirooka, A. Schwarz *et al.*, 2017 Description of two three-gendered nematode species in the new genus *Auanema* (Rhabditina) that are models for reproductive mode evolution. *Sci Rep* 7 (1):11135.

Kiontke, K., A. Barriere, I. Kolotuev, B. Podbilewicz, R. Sommer *et al.*, 2007 Trends, stasis, and drift in the evolution of nematode vulva development. *Curr Biol* 17 (22):1925-1937.

Kiontke, K., N.P. Gavin, Y. Raynes, C. Roehrig, F. Piano *et al.*, 2004 *Caenorhabditis* phylogeny predicts convergence of hermaphroditism and extensive intron loss. *Proc Natl Acad Sci U S A* 101 (24):9003-9008.

Korf, I., 2004 Gene finding in novel genomes. *BMC Bioinformatics* 5:59.

Krzywinski, M., J. Schein, I. Birol, J. Connors, R. Gascoyne *et al.*, 2009 Circos: an information aesthetic for comparative genomics. *Genome Res* 19 (9):1639-1645.

Kumar, S., M. Jones, G. Koutsovoulos, M. Clarke, and M. Blaxter, 2013 Blobology: exploring raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-coverage plots. *Front Genet* 4:237.

Kurtz, S., A. Phillippy, A.L. Delcher, M. Smoot, M. Shumway *et al.*, 2004 Versatile and open software for comparing large genomes. *Genome Biol* 5 (2):R12.

Lande, R., and D.W. Schemske, 1985 The evolution of self-fertilization and inbreeding depression in plants .1. Genetic Models. *Evolution* 39 (1):24-40.

Li, H., 2011 A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27 (21):2987-2993.

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25 (14):1754-1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25 (16):2078-2079.

Liao, Y., G.K. Smyth, and W. Shi, 2014 featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30 (7):923-930.

Loewe, L., and A.D. Cutter, 2008 On the potential for extinction by Muller's ratchet in *Caenorhabditis elegans*. *BMC Evol Biol* 8:125.

Love, M.I., W. Huber, and S. Anders, 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15 (12):550.

Lowe, T.M., and S.R. Eddy, 1997 tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25 (5):955-964.

Luo, R., B. Liu, Y. Xie, Z. Li, W. Huang *et al.*, 2012 SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1 (1):18.

Margarido, G.R., A.P. Souza, and A.A. Garcia, 2007 OneMap: software for genetic mapping in outcrossing species. *Hereditas* 144 (3):78-79.

Maupas, E., 1900 Modes et formes de reproduction des nématodes. *Ann Zool Exp Gen* 8 (3):463-624.

Maynard Smith, J., 1978 *The evolution of sex*. Cambridge Cambridge University Press.

Milne, I., G. Stephen, M. Bayer, P.J. Cock, L. Pritchard *et al.*, 2013 Using Tablet for visual exploration of second-generation sequencing data. *Brief Bioinform* 14 (2):193-202.

Minniti, A.N., C. Sadler, and S. Ward, 1996 Genetic and molecular analysis of *spe-27*, a gene required for spermiogenesis in *Caenorhabditis elegans* hermaphrodites. *Genetics* 143 (1):213-223.

Myhre, S., H. Tveit, T. Mollestad, and A. Laegreid, 2006 Additional gene ontology structure for improved biological reasoning. *Bioinformatics* 22 (16):2020-2027.

Nawrocki, E.P., and S.R. Eddy, 2013 Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29 (22):2933-2935.

Nigon, V.M., and M.A. Felix, 2017 History of research on *C. elegans* and other free-living nematodes as model organisms. *WormBook*:1-91.

Otto, S.P., C. Sassaman, and M.W. Feldman, 1993 Evolution of sex determination in the conchostracan shrimp *Eulimnadia texana*. *Am Nat* 141 (2):329-337.

1081 Parra, G., K. Bradnam, and I. Korf, 2007 CEGMA: a pipeline to accurately annotate core
1082 genes in eukaryotic genomes. *Bioinformatics* 23 (9):1061-1067.
1083 Perrin, N., 2016 Random sex determination: When developmental noise tips the sex
1084 balance. *Bioessays* 38 (12):1218-1226.
1085 Pires-daSilva, A., 2007 Evolution of the control of sexual identity in nematodes. *Semin*
1086 *Cell Dev Biol* 18 (3):362-370.
1087 Pires-daSilva, A., 2013 *Pristionchus pacificus* protocols. *WormBook*:1-20.
1088 Pires-daSilva, A., and R.J. Sommer, 2004 Conservation of the global sex determination
1089 gene *tra-1* in distantly related nematodes. *Genes Dev* 18 (10):1198-1208.
1090 Poinar, G.O., 1983 *The natural history of nematodes*. Englewood Cliffs, N.J.: Prentice-
1091 Hall.
1092 Quinlan, A.R., and I.M. Hall, 2010 BEDTools: a flexible suite of utilities for comparing
1093 genomic features. *Bioinformatics* 26 (6):841-842.
1094 Rödelsperger, C., J.M. Meyer, N. Prabh, C. Lanz, F. Bemm *et al.*, 2017 Single-molecule
1095 sequencing reveals the chromosome-scale genomic architecture of the
1096 nematode model organism *Pristionchus pacificus*. *Cell Reports* 21 (3):834-844.
1097 Schedl, T., and J. Kimble, 1988 *fog-2*, a germ-line-specific sex determination gene
1098 required for hermaphrodite spermatogenesis in *Caenorhabditis elegans*.
1099 *Genetics* 119 (1):43-61.
1100 Schulz, M.H., D. Weese, M. Holtgrewe, V. Dimitrova, S. Niu *et al.*, 2014 Fiona: a parallel
1101 and automatic strategy for read error correction. *Bioinformatics* 30 (17):i356-363.
1102 Shakes, D.C., B.J. Neva, H. Huynh, J. Chaudhuri, and A. Pires-daSilva, 2011
1103 Asymmetric spermatocyte division as a mechanism for controlling sex ratios. *Nat*
1104 *Commun* 2:157.
1105 Shen, Y., and R.E. Ellis, 2018 Reproduction: sperm with two X chromosomes and eggs
1106 with none. *Curr Biol* 28 (3):R121-R124.
1107 Smit, A.F.A., and R. Hubley, 2008-2015 RepeatModeler Open-1.0.
1108 Smit, A.F.A., and R. Hubley, 2013-2015 RepeatMasker-4.0.
1109 Stanke, M., and S. Waack, 2003 Gene prediction with a hidden Markov model and a
1110 new intron submodel. *Bioinformatics* 19 Suppl 2:ii215-225.

Stein, L.D., Z. Bao, D. Blasiar, T. Blumenthal, M.R. Brent *et al.*, 2003 The Genome Sequence of *Caenorhabditis briggsae*: A Platform for Comparative Genomics. *PLoS Biol* 1 (2):E45.

Stewart, A.D., and P.C. Phillips, 2002 Selection and maintenance of androdioecy in *Caenorhabditis elegans*. *Genetics* 160 (3):975-982.

Stiernagle, T., 2006 Maintenance of *C. elegans*. *WormBook*:1-11.

Strome, S., W.G. Kelly, S. Ercan, and J.D. Lieb, 2014 Regulation of the X chromosomes in *Caenorhabditis elegans*. *Cold Spring Harb Perspect Biol* 6 (3):a018366.

Sudhaus, W., 1976 Vergleichende Untersuchungen zur Phylogenie, Systematik, Ökologie und Ethologie der Rhabditidae (Nematoda). *Zoologica* 43:1-229.

Sudhaus, W., and D.H. Fitch, 2001 Comparative studies on the phylogeny and systematics of Rhabditidae (Nematoda). *J Nematol* 33 (1):1-72.

Tandonnet, S., M.C. Farrell, G.D. Koutsovoulos, M.L. Blaxter, M. Parihar *et al.*, 2018 Sex- and gamete-specific patterns of X chromosome segregation in a trioecious nematode. *Curr Biol* 28 (1):93-99 e93.

Ter-Hovhannisyan, V., A. Lomsadze, Y.O. Chernoff, and M. Borodovsky, 2008 Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res* 18 (12):1979-1990.

Triantaphyllou, A.C., and H. Hirschmann, 1964 Reproduction in plant and soil nematodes. *Annu Rev Phytopathol* 2:57-80.

Walton, A.C., 1959 Some parasites and their chromosomes. *J Parasitol* 45:1-20.

Weeks, S.C., 2012 The role of androdioecy and gynodioecy in mediating evolutionary transitions between dioecy and hermaphroditism in the Animalia. *Evolution* 66 (12):3670-3686.

Weeks, S.C., C. Benvenuto, and S.K. Reed, 2006a When males and hermaphrodites coexist: a review of androdioecy in animals. *Integr Comp Biol* 46 (4):449-464.

Weeks, S.C., B.R. Crosser, R. Bennett, M. Gray, and N. Zucker, 2000 Maintenance of androdioecy in the freshwater shrimp, *Eulimnadia texana*: estimates of inbreeding depression in two populations. *Evolution Int J Org Evolution* 54 (3):878-887.

Weeks, S.C., T.F. Sanderson, S.K. Reed, M. Zofkova, B. Knott *et al.*, 2006b Ancient androdioecy in the freshwater crustacean *Eulimnadia*. *Proc Biol Sci* 273 (1587):725-734.

Wickham, H., 2016 *Ggplot2*. New York, NY: Springer Science+Business Media, LLC.

Wilson, R.K., 1999 How the worm was won: the *C. elegans* genome sequencing project. *Trends in Genetics* 15 (2):51-58.

Wright, S.I., R.W. Ness, J.P. Foxe, and S.C.H. Barrett, 2008 Genomic consequences of outcrossing and selfing in plants. *Int J Plant Sci* 169 (1):105-118.

Xiong, Y., X. Chen, Z. Chen, X. Wang, S. Shi *et al.*, 2010 RNA sequencing shows no dosage compensation of the active X-chromosome. *Nat Genet*.

Yin, D., E.M. Schwarz, C.G. Thomas, R.L. Felde, I.F. Korf *et al.*, 2018 Rapid genome shrinkage in a self-fertile nematode reveals sperm competition proteins. *Science* 359 (6371):55-61.

Zarkower, D., 2013 DMRT genes in vertebrate gametogenesis. *Curr Top Dev Biol* 102:327-356.

Zarkower, D., and J. Hodgkin, 1992 Molecular analysis of the *C. elegans* sex-determining gene *tra-1*: a gene encoding two zinc finger proteins. *Cell* 70 (2):237-249.

Zioni (Cohen-Nissan), S., I. Glazer, and D. Segal, 1992 Life-cycle and reproductive potential of the nematode *Heterorhabditis bacteriophora* strain Hp88. *J Nematol* 24 (3):352-358.

Figures

Figure 1. X chromosome segregation and types of progeny produced in each type of cross in *A. rhodensis*. Due to the unique meiosis of the X chromosome, hermaphrodites produce mostly XX progeny (hermaphrodites and females) by selfing

(**A**) and XO male cross-progeny (**B**). Males produce only one type of functional sperm, haplo-X (**B** and **C**), although they are XO. Smaller cells denotes rarer events. Nullo-X, haplo-X and diplo-X gametes are indicated by 0X, 1X and 2X, respectively. Colors refer to egg (orange) and sperm (blue) contributions.

Figure 2. The relationship of *Auanema rhodensis* to other rhabditid nematodes.

The phylogeny of the nematode species discussed in this analysis. *A. rhodensis* and *O. tipulae* are sister taxa in analyses based on multiple protein coding gene and ribosomal RNA loci (Blaxter and Koutsovoulos 2014; Kanzaki et al. 2017; Kiontke et al. 2007).

Figure 3. Synteny relationships between chromosomes of *Auanema rhodensis* and *Caenorhabditis elegans*.

Location (**A** and **C**) and proportion (**B** and **D**) of orthologous protein-coding genes between *C. elegans* and *A. rhodensis* coloured according to *A. rhodensis* (**A** and **B**) or *C. elegans* (**C** and **D**) chromosomes. Each line in the circos plots corresponds to a predicted orthologous gene between the two species.

Figure 4. Macrosynteny relationships of *Auanema rhodensis* linkage groups LG2, LG3 and LG4.

The circos plots show macrosynteny relationships of (columns) LG2 (yellow), LG3 (pink) and LG4 (blue) of *A. rhodensis* to (rows) the chromosomal genomes of C. elegans, *H.*

contortus, O. tipulae and *P. pacificus*, based on the mapping of presumed orthologues between the species. Each line in the circos plots corresponds to a predicted orthologous gene between the two species.

Figure 5. Macrosynteny relationships of rhabditine X chromosomes.

For four rhabditine nematodes for which chromosomal genome assemblies or scaffold allocations to chromosomes (*O. tipulae*) are available, we mapped the location of their X-linked genes to the *A. rhodensis* genome (**A**: *Pristionchus pacificus*; **B**: *O. tipulae*; **C**: *Haemonchus contortus*; **D**: *Caenorhabditis elegans*). **E** Distribution of mappings to *A. rhodensis* chromosomes of X-linked genes in *H. contortus* (upper) and *C. elegans* (lower). The X of *O. tipulae* is represented as a concatenation of all the scaffolds belonging to the X chromosome; their order is arbitrary (scaffold number).

Figure 6. Contrasting genomic patterns between the X chromosome and the autosomes.

(A) Distribution and conservation of protein-coding genes across *A. rhodensis* chromosomes. Density of *A. rhodensis* protein coding genes (upper panel) and conserved genes between *A. rhodensis* and *D. melanogaster* (lower panel) along each linkage group using a 200,000 bp window size. Overall, 4,544 conserved genes were identified between *A. rhodensis* and *D. melanogaster*. **(B)** Localisation of the annotated non-coding genes (upper panel) and transfer RNAs (tRNAs) (lower panel) along each linkage group using a 300,000 bp window size. No tRNAs were found on the X

chromosome (LG5). **(C)** Patterns of variation across two inbred strains of *A. rhodensis*. Variant density along each chromosome in 250,000 base windows for the within-strain variants (upper panels) and 100,000 base windows for the between-strain variants (lower panel). **(D)** Genotype frequencies across *A. rhodensis* chromosomes. Black and red lines represent the frequencies of RAD sites homozygous for the APS4 or for the APS6 allele, respectively, in the 95 genotyped F2Ls. Blue lines represent heterozygous genotypes. More than 80% of the progeny samples were heterozygous for the X chromosome (LG5X). The black ticks on the x-axis show the positions of the 1052 mapped RAD markers.

Figure 7. Expression of genes on the *A. rhodensis* X chromosome is generally lower than those on autosomes.

Boxplots of the log2 normalized expression of the genes located on each linkage group of *A. rhodensis* in different sexes and stages in single replicate libraries. The expression levels were normalized by library size and log2-transformed. LG5 (in red) is the X chromosome. Boxplots for all libraries are represented in Figure S4. This plot was generated using the R package ggplot2 (Wickham 2016).

Figure 8. Analysis of differential gene expression in *Auanema rhodensis* XX nematodes.

(A) Differentially expressed genes in *A. rhodensis* XX L2 larvae. Most genes found to be DE between female L2 and hermaphrodite L2 were also differentially expressed

between the converted female L2 and the hermaphrodite L2. Few DE genes were found between the female L2 and converted female L2. Of these, 32 had similar expression in hermaphrodite L2. **(B)** Distribution of DE genes (absolute log2(FC) >= 2, FDR <0.01) along the chromosomes of *A. rhodensis*. Upper panel: Female L2 *versus* converted female L2; middle panel: hermaphrodite L2 *versus* converted female L2; lower panel: female L2 *versus* hermaphrodite L2. A plot of the density of all genes along the genome (see Figure 6C) is shown under the lower panel.

Figure 9: Nigon elements and the evolution of rhabditine chromosomes.

The different colors indicate orthologous chromosomes/chromosomal parts belonging to different Nigon elements. Nigon element '*NN*' may be part of NX or NE or a separate unit as depicted here. Reshuffling within chromosomes is not depicted.

Table 1. Genome and transcriptome libraries used in this study*

Type	***Strain***	***SRA Accession***	***Sample name***	***Library type******	***Raw read pairs***
Genome	APS4	ERS3048742 (SAMEA5241338)	APS4_250bp_1	PE	30115881

		ERS3048743 (SAMEA5241339)	APS4_250bp_2	PE	44755823
		ERS3048744 (SAMEA5241340)	APS4_250bp_3	PE	48130281
	APS4	ERS3048745 (SAMEA5241341)	APS4_3kb	MP	141654936
		ERS3048746 (SAMEA5241342)	APS4_6kb	MP	108976027
	APS6	ERS3048747 (SAMEA5241343)	APS6_450bp	PE	24219024
Genome (RAD-seq)	F2Ls between APS4 and APS6	ERS3048748 (SAMEA5241344)	2014132_MBlib1 (24 samples)	RAD	44063445
		ERS3048749 (SAMEA5241345)	2014132_MBlib2 (24 samples)	RAD	47384155
		ERS3048750 (SAMEA5241346)	2014132_MBlib3 (24 samples)	RAD	48905114
		ERS3048751 (SAMEA52413	2014132_MBlib4 (25 samples)	RAD	39742808

		47)			
Transcriptome	APS4	See **	L2_fem_lib1	RNA	17110716
			L2_fem_lib2	RNA	16629611
			L2_fem_lib3	RNA	17379062
			L2_DA_fem_lib1	RNA	18536076
			L2_DA_fem_lib2	RNA	17170105
			L2_DA_fem_lib3	RNA	16962402
			L2_herm_lib1	RNA	15562977
			L2_herm_lib2	RNA	16078424
			L2_herm_lib3	RNA	16426418
			Males_lib1	RNA	16341785
			Males_lib2	RNA	16760682
			Males_lib3	RNA	13624819
			Mixed_stages_lib1	RNA	24190447
			Mixed_stages_lib2	RNA	22250315
			Mixed_stages_lib3	RNA	33719128

1250

1251 * All genomic data have been submitted under Bioproject number PRJEB29492

1252 ** RNA-seq data have been deposited in the ArrayExpress database at EMBL-EBI

1253 (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-7667.

1254 *** PE: paired-end genomic; MP: mate-pair genomic; RAD: paired-end RAD-seq; RNA:

1255 paired-end RNA-seq

Table 2. Genome assembly metrics of *Auanema rhodensis*, *Caenorhabditis elegans* and *O. tipulae*.

Metric	***A. rhodensis* (scaffolds)**	***A. rhodensis* (chromosomal assembly)**	***C. elegans* (PRJNA13758)***	***O. tipulae*** (CEW1_nOt2)
Number of scaffolds or chromosomes	636	6 autosomes 1 X 493 unplaced scaffolds	5 autosomes 1 X	191
Assembly size or draft genome size (Mb)	60.6	57.8	100.2	59.4
Number of scaffolds (> 200 bp)	636	7	6	191
Number of scaffolds (> 1,000 bp)	440	7	6	157
N50 (scaffolds > 1,000 bp) (bp)	556,081	8,804,062	17,493,829	1,203,411
Longest scaffold / chromosome (bp)	3,360,731	9,627,060	20,924,180	4,597,891
GC content (%)	32.2	32.2	35.44	44.53
Span of runs of Ns (>= 10 Ns) (bp)	915,180	928,780	NA	16310
Protein coding gene annotations				
Number of genes (protein-coding)	11,570	10,861	23,629	14,938
Exons				
Number of coding exons	135,144	130,644	189,079	127,820
Combined length of exons (bp)	16,655,465	15,869,469	39,400,137	20,438,569
Exon mean length (SD) (bp)	123.2 (131.1)	121.5 (121.8)	208.4 (263)	159.9 (138.09)

Exon median length (bp)	109	109	146	132
Minimum/maximum exon length (bp)	3 / 11,659	3 / 11,659	1 / 14,975	3 / 10,071
Introns				
Number of introns	123,693	119,812	200,020	112,925
Combined length of introns (bp)	16,862,509	16,308,478	79,153,867	17,941,595
Intron mean length (SD) (bp)	136.3 (521.3)	136.1 (521.4)	395.7 (962.3)	158.9 (366.8)
Intron median length (bp)	47	47	82	41
Minimum/maximum intron length (bp)	7 / 24,877	7 / 24,877	1 / 100,912	7 / 11,345

***** Metrics for *C. elegans* were calculated from the WormBase annotations (WormBase web site, http://www.wormbase.org, release WS262, July 2018).

Table 3. Characteristics of the *A. rhodensis* genetic map.

Linkage group	Number of RAD-seq markers	Homozygous APS4/APS4 frequency	Heterozygous APS4/APS6 frequency	Homozygous APS6/APS6 frequency	Assembly length (bp) *	Number of protein-coding genes
LG1	109	0.30	0.35	0.35	8,489,927	1,538
LG2	149	0.29	0.37	0.34	9,627,060	1,760
LG3	184	0.21	0.38	0.51	8,741,542	1,748
LG4	143	0.35	0.39	0.36	8,804,062	1,586
LG5X	92	0.04	0.88	0.08	3,488,253	604

LG6	185	0.35	0.39	0.26	9,421,540	1,871
LG7	190	0.23	0.42	0.35	9,306,279	1,754
Overall	1052	0.26	0.42	0.32	57,878,663	10,861**

1262

1263 * Length in chromosomal assembly after anchoring the scaffolds onto the genetic map.

1264 ** 93% of total number of protein coding genes.

1265

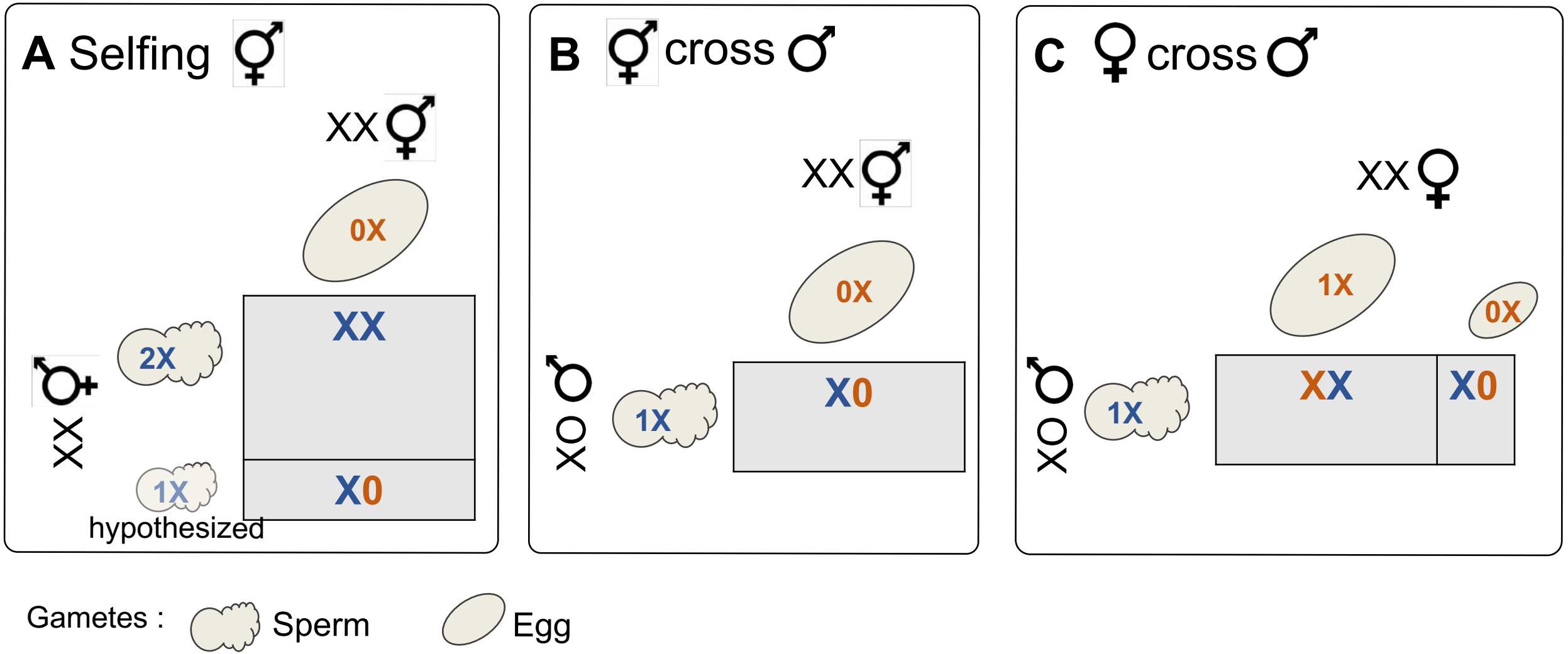
1266 **Table 4. Inhibition of the DM transcription factor Arh-g5747 (dmd-10/11) by RNAi**

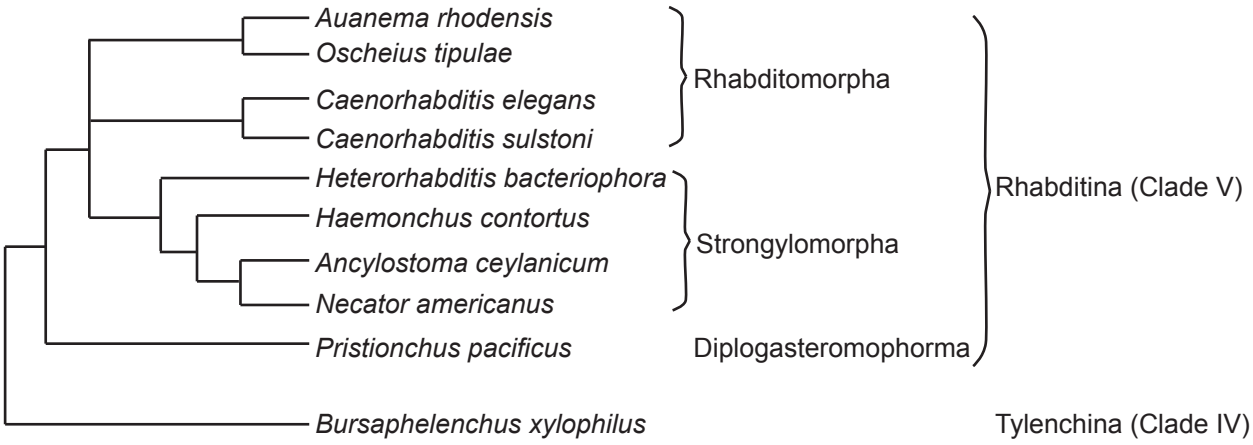
1267 **in hermaphrodite mothers results in more female progeny.**

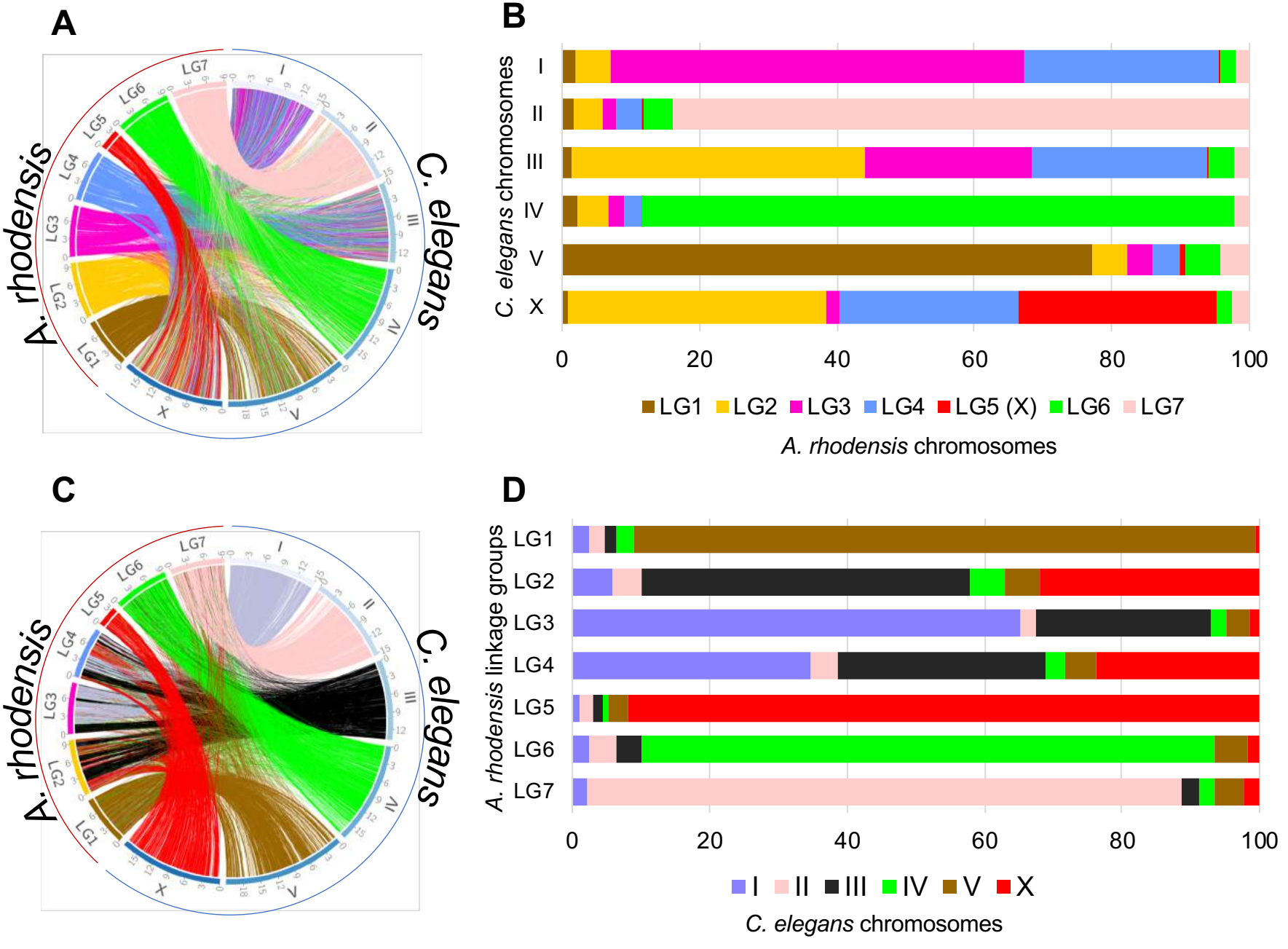
Condition	**Hermaphrodite Injected**	**Progeny of injected animals**			
		Males	**Females**	**Hermaphrodites**	**Female Ratio [Females / (Females + Hermaphrodites)]**
RNAi	1	21	57	61	0.48
RNAi	2	12	90	76	0.54
RNAi	3	7	82	185	0.31
RNAi	4	18	101	227	0.31
RNAi	5	13	94	52	0.64
RNAi	6	2	95	89	0.52
RNAi	7	4	55	88	0.38
RNAi	8	2	53	64	0.45

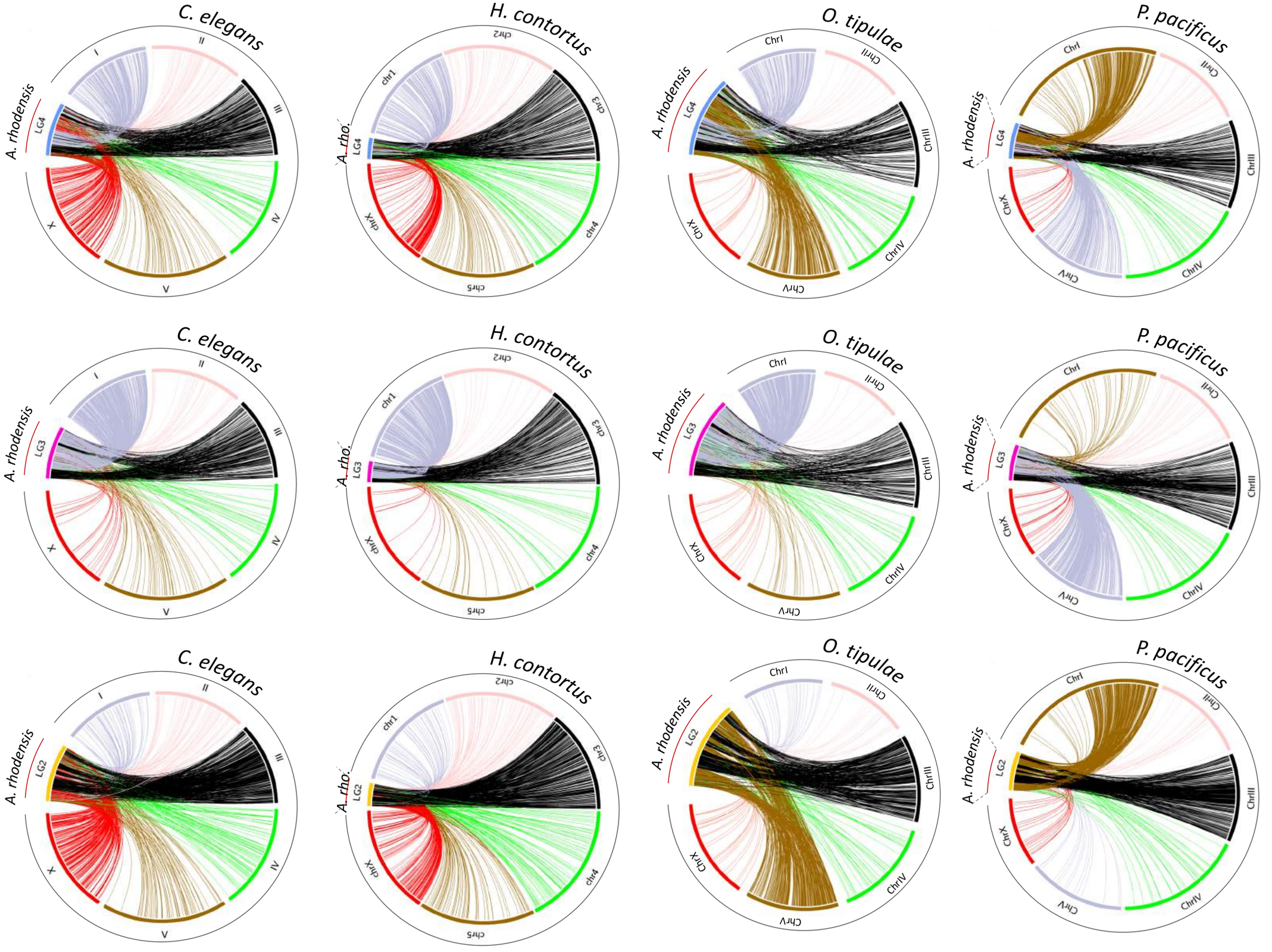
RNAi	***8 mothers***	***9.88***	***78.38***	***105.25***	***0.43***
Control	9	3	33	107	0.24
Control	10	4	43	254	0.14
Control	11	12	44	329	0.12
Control	12	1	4	20	0.17
Control	13	0	26	40	0.39
Control	14	15	53	158	0.25
Control	15	3	25	70	0.26
Control	16	12	75	190	0.28
Control	17	8	69	142	0.33
Control	***9 mothers***	***6.44***	***41.33***	***145.56***	***0.22***

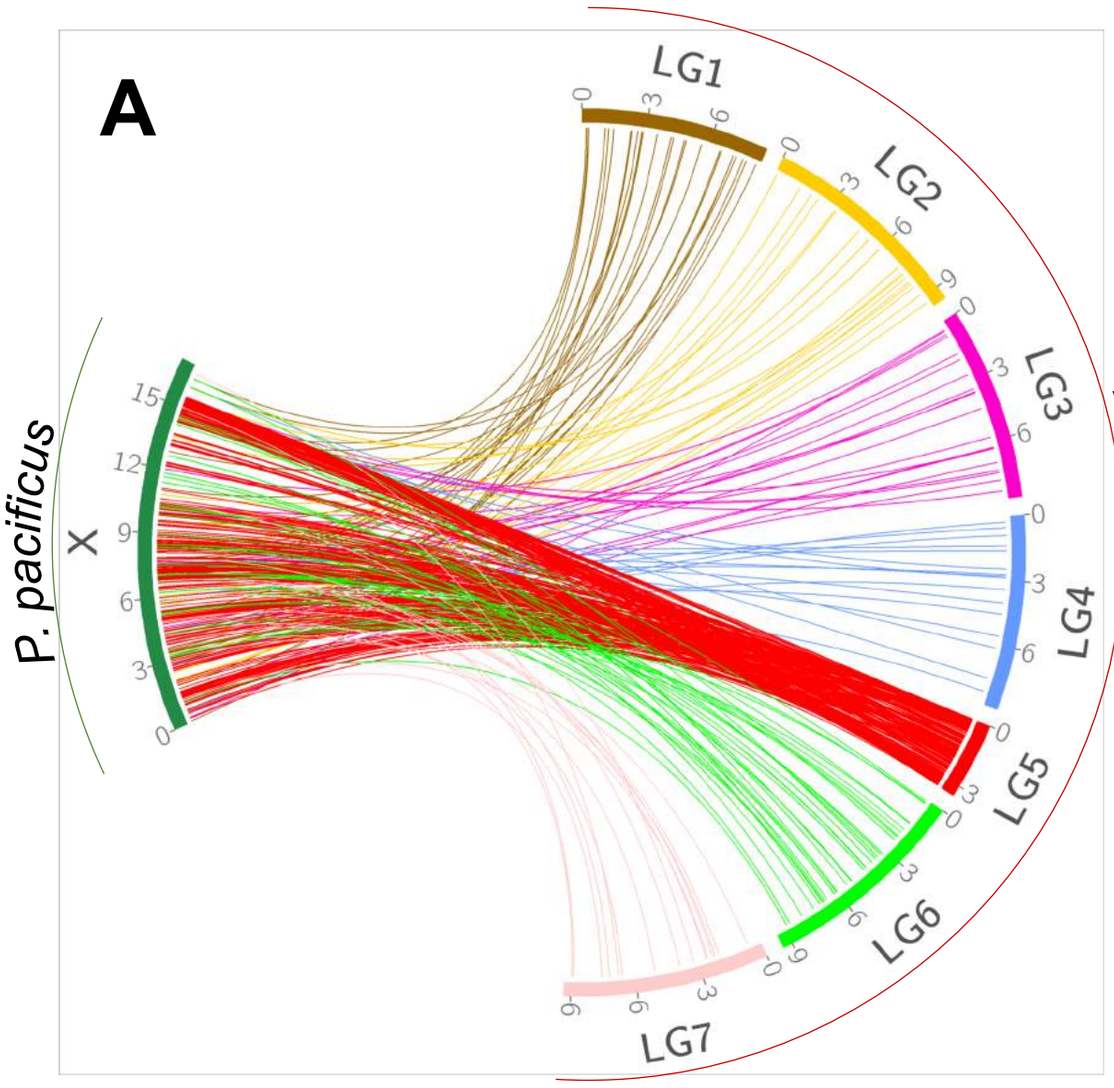
1268

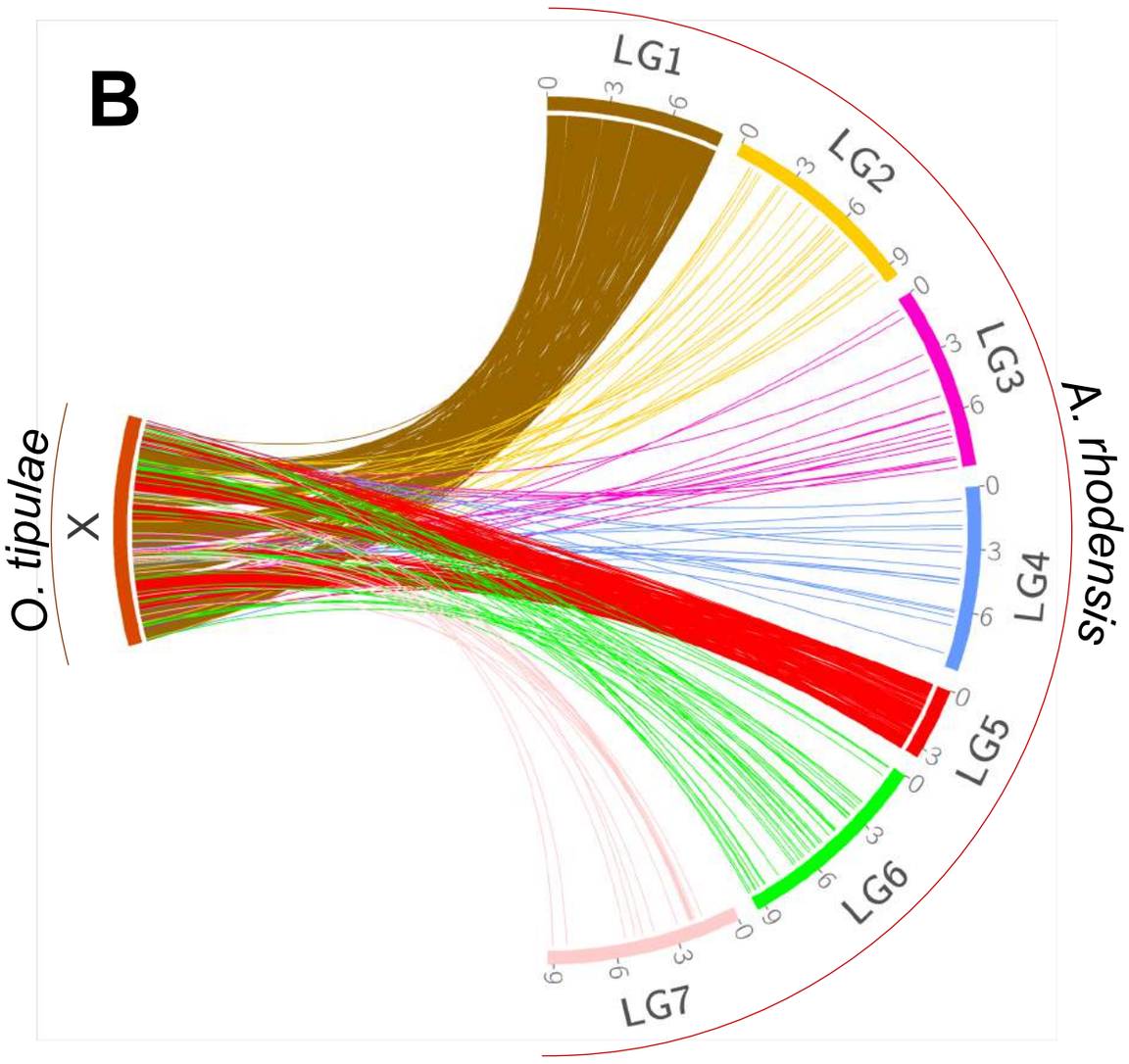

A Selfing
XX
0X
XX
2X
1X
X0
hypothesized
B cross
XX
0X
XO
1X
X0
C cross
XX
1X
0X
XO
1X
XX
X0
Gametes :
Sperm
Egg

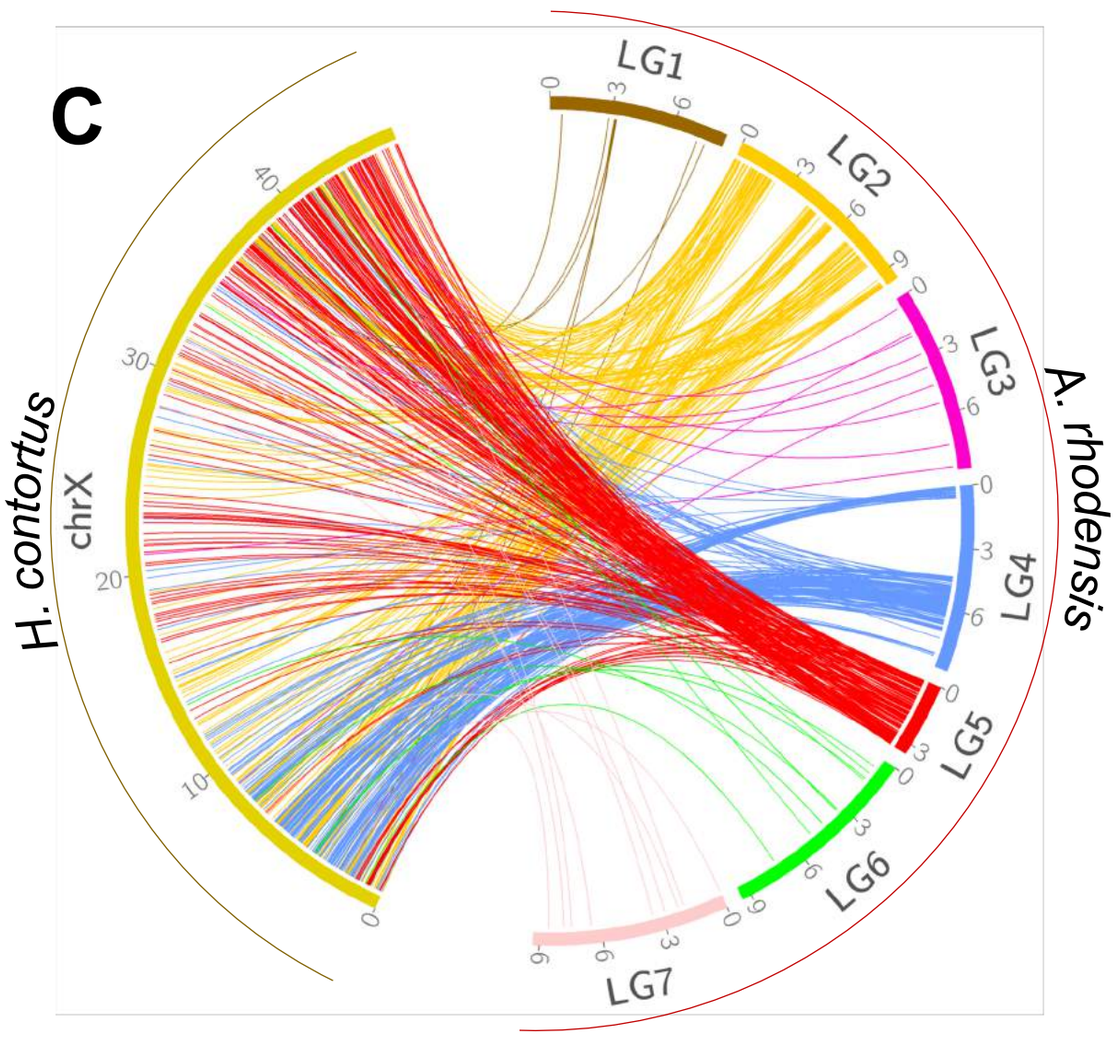

Auanema rhodensis
Oscheius tipulae
Caenorhabditis elegans
Caenorhabditis sulstoni
Heterorhabditis bacteriophora
Haemonchus contortus
Ancylostoma ceylanicum
Necator americanus
Pristionchus pacificus
Bursaphelenchus xylophilus
Rhabditomorpha
Strongylomorpha
Diplogasteromophorma
Rhabditina (Clade V)
Tylenchina (Clade IV)

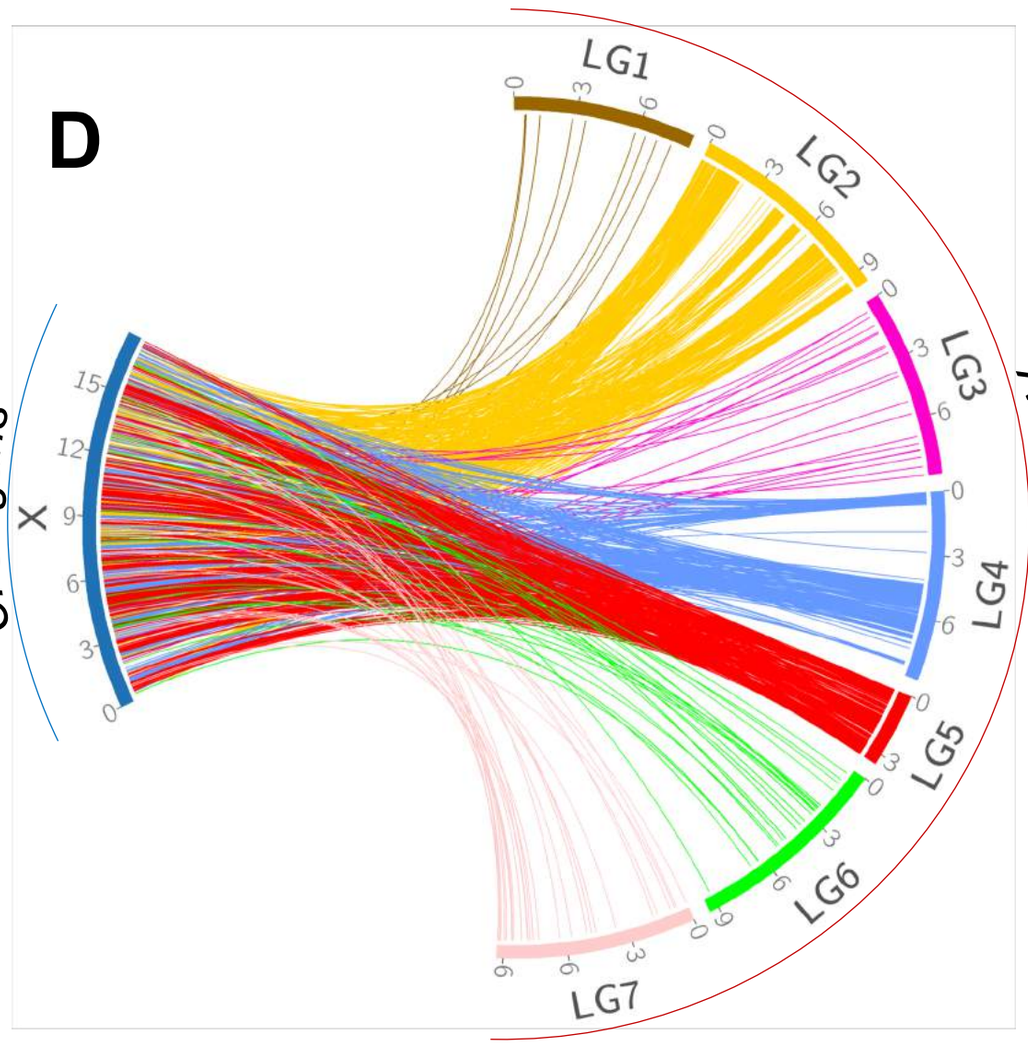

A
A. rhodensis
C. elegans
LG1
LG2
LG3
LG4
LG5
LG6
LG7
I
II
III
IV
V
X
B
C. elegans chromosomes
I
II
III
IV
V
X
0
20
40
60
80
100
LG1 LG2 LG3 LG4 LG5 (X) LG6 LG7
A. rhodensis chromosomes
C
A. rhodensis
C. elegans
D
A. rhodensis linkage groups
LG1
LG2
LG3
LG4
LG5
LG6
LG7
0
20
40
60
80
100
I II III IV V X
C. elegans chromosomes

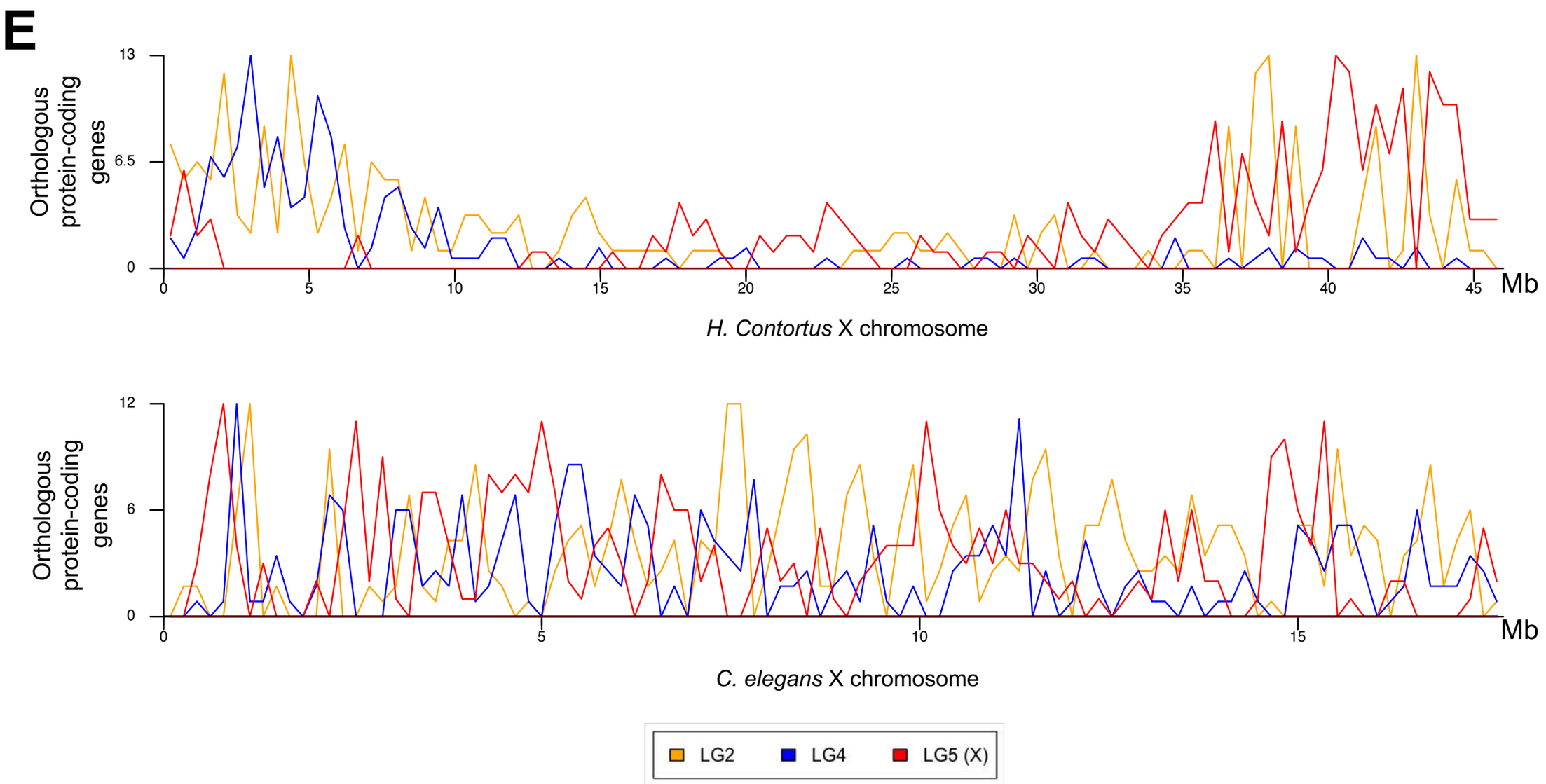

C. elegans
H. contortus
O. tipulae
P. pacificus
A. rhodensis
LG4
LG3
LG2
I
II
III
IV
V
X
chr1
chr2
chr3
chr4
chr5
chrX
ChrI
ChrII
ChrIII
ChrIV
ChrV
ChrX

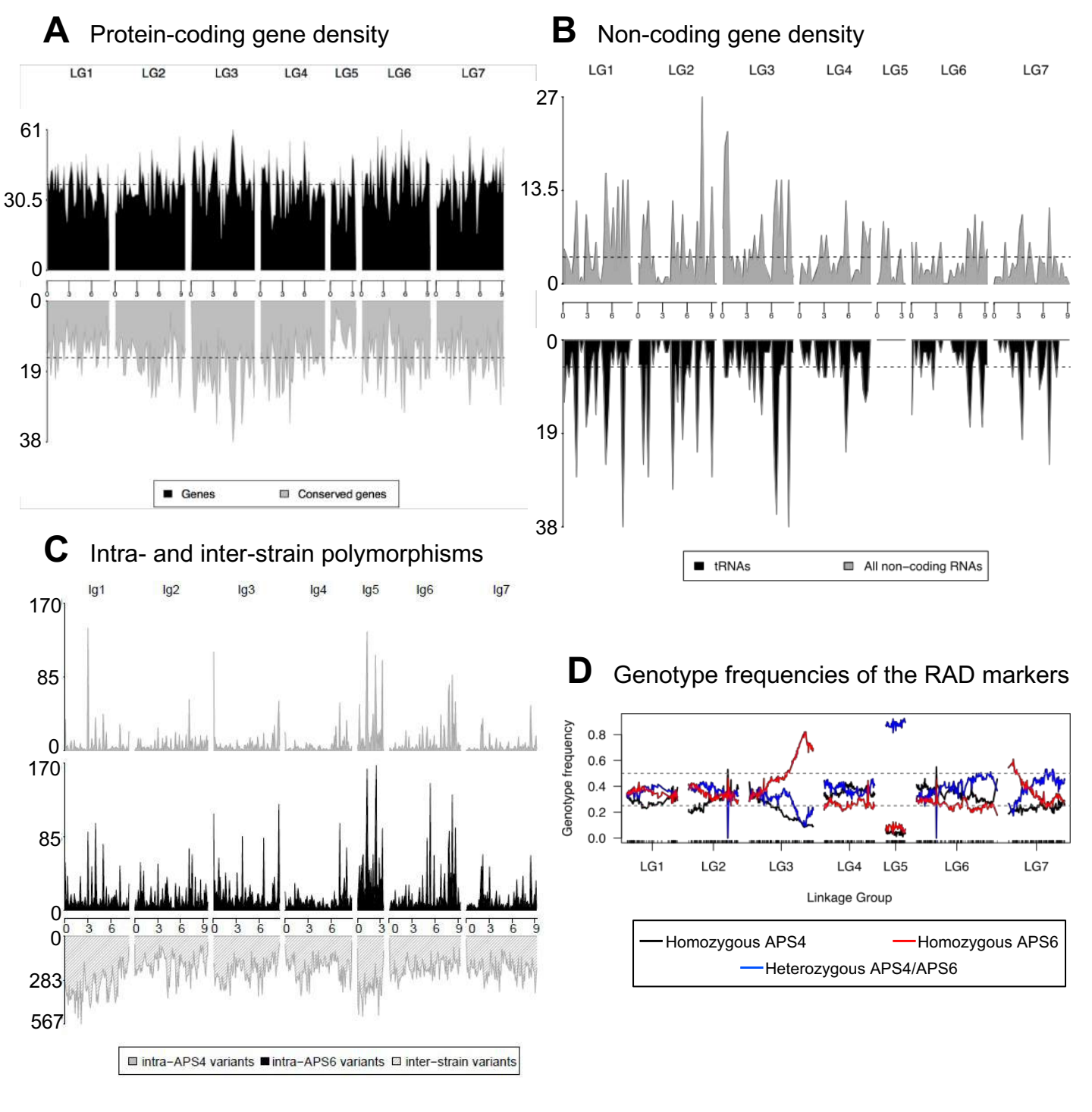

A
P. pacificus
A. rhodensis
X
LG1
LG2
LG3
LG4
LG5
LG6
LG7

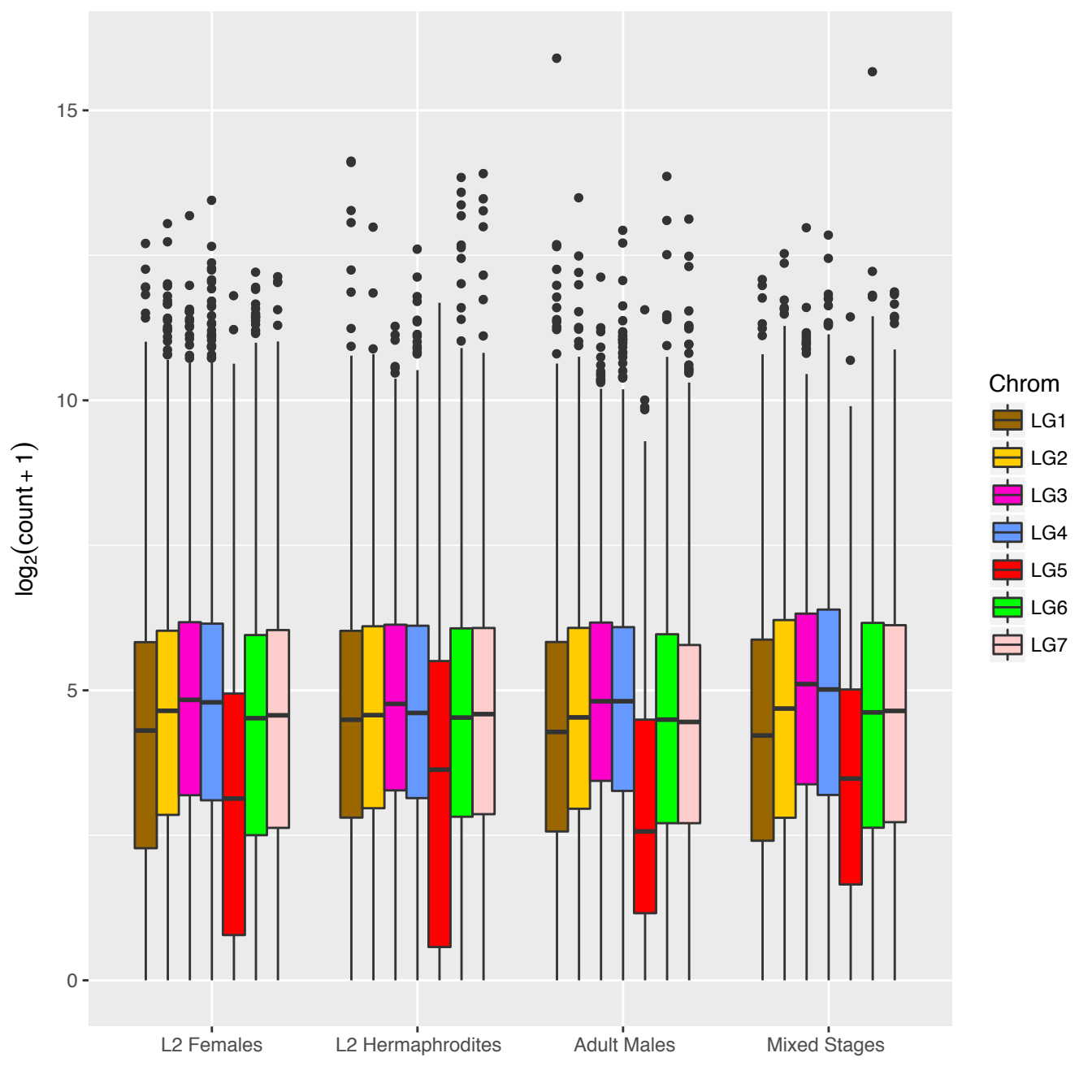

B
O. tipulae
A. rhodensis
X
LG1
LG2
LG3
LG4
LG5
LG6
LG7

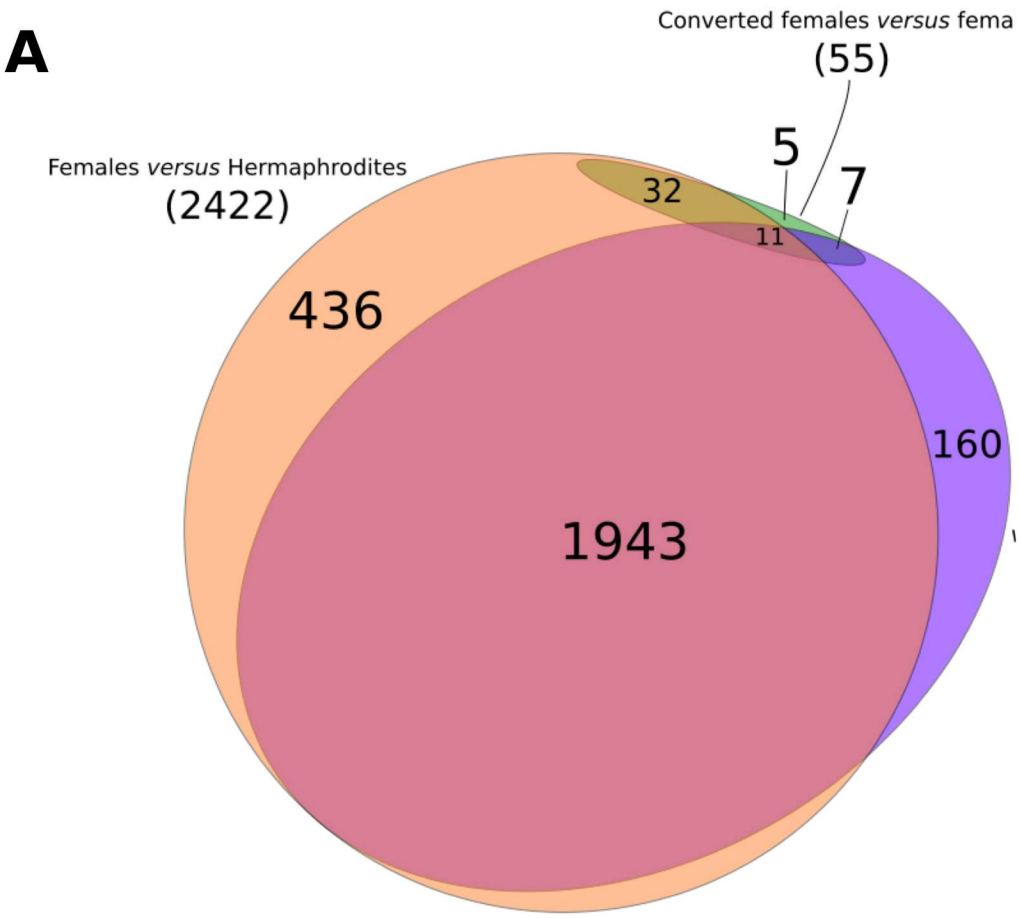

C
H. contortus
A. rhodensis
chrX
LG1
LG2
LG3
LG4
LG5
LG6
LG7

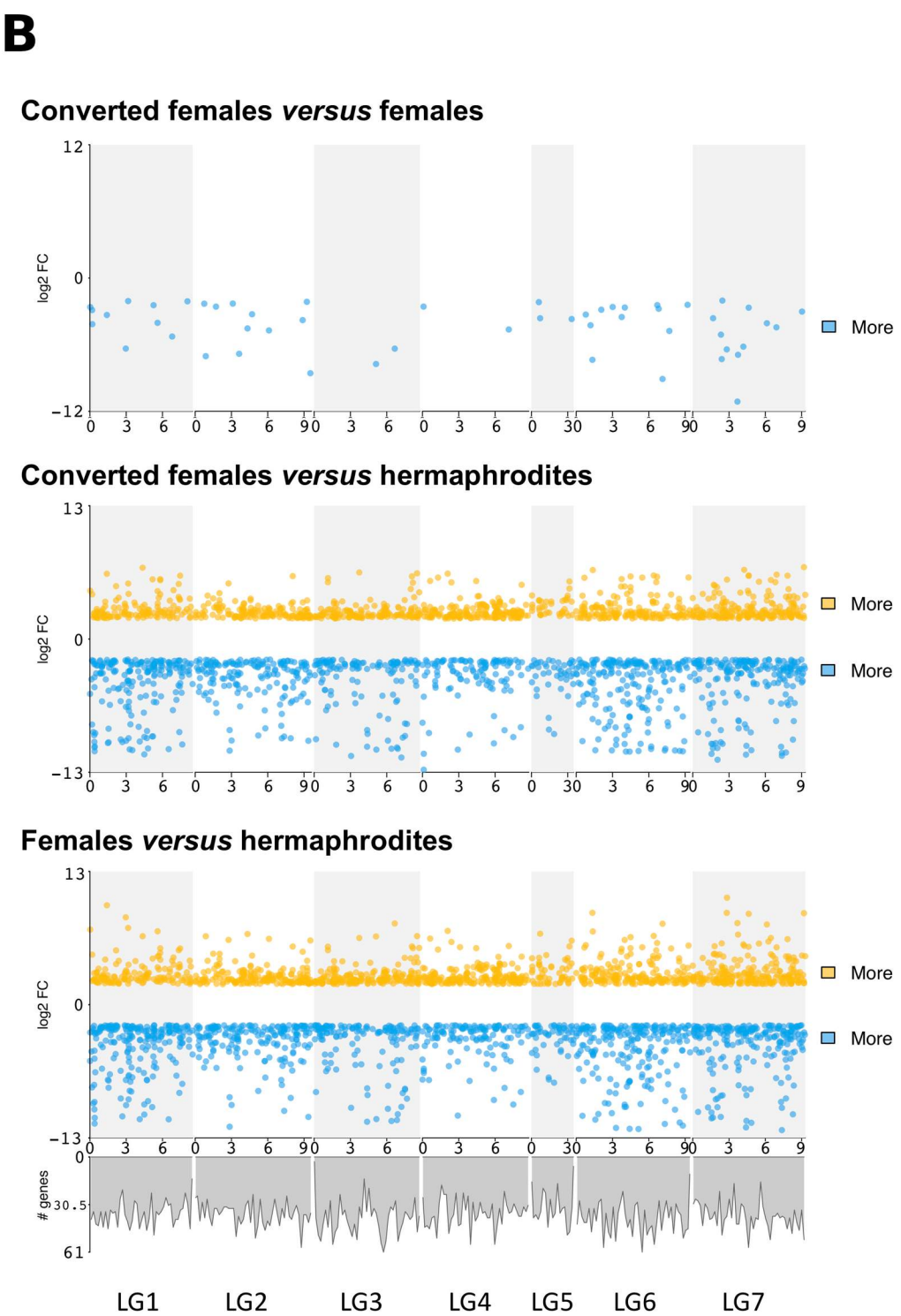

D
C. elegans
A. rhodensis
X
LG1
LG2
LG3
LG4
LG5
LG6
LG7

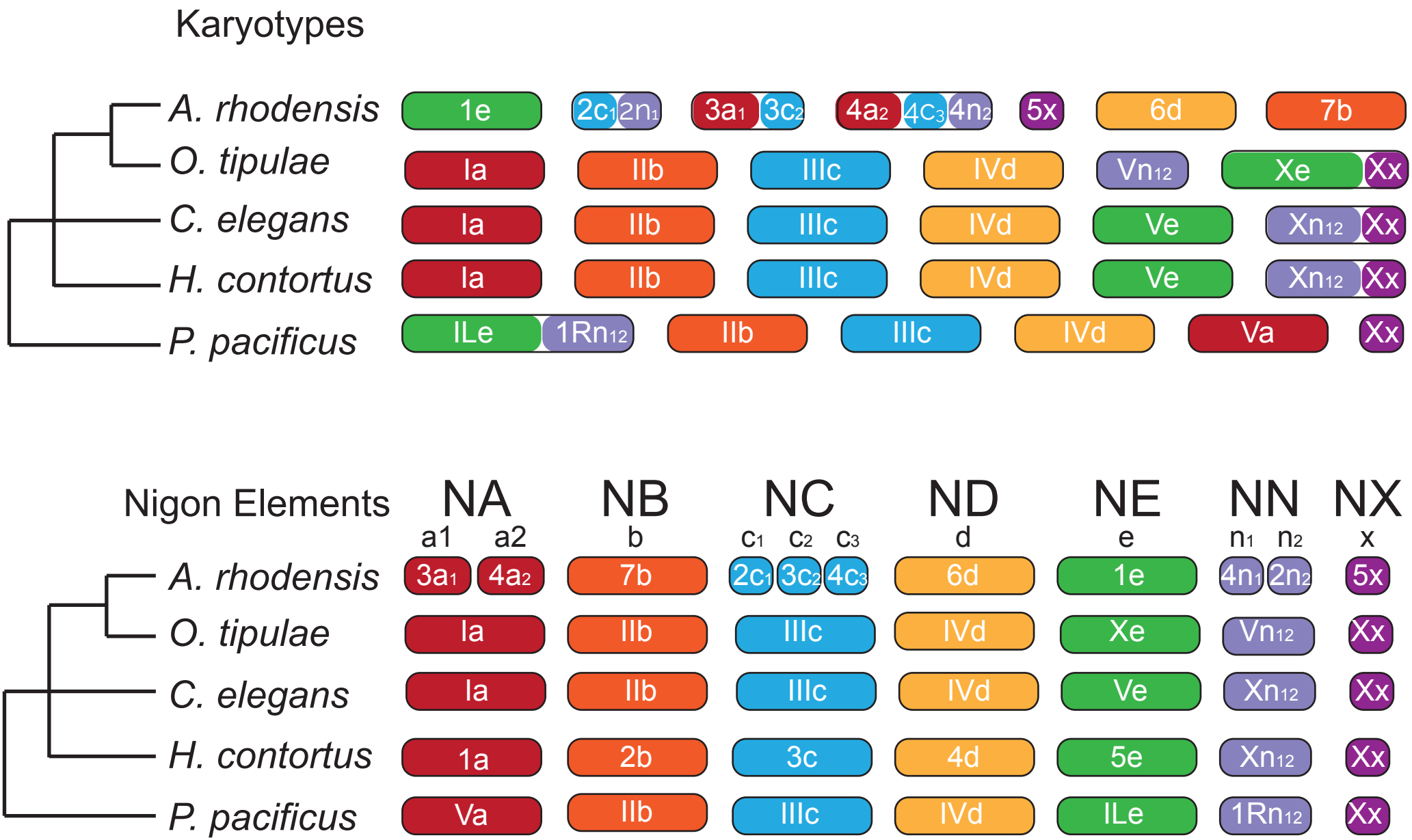

E
Orthologous protein-coding genes
13
6.5
0
0
5
10
15
20
25
30
35
40
45
Mb
H. Contortus X chromosome
Orthologous protein-coding genes
12
6
0
0
5
10
15
Mb
C. elegans X chromosome
LG2
LG4
LG5 (X)


A
Protein-coding gene density
LG1
LG2
LG3
LG4
LG5
LG6
LG7
61
30.5
0
0
19
38
Genes
Conserved genes
B
Non-coding gene density
LG1
LG2
LG3
LG4
LG5
LG6
LG7
27
13.5
0
0
19
38
tRNAs
All non−coding RNAs
C
Intra- and inter-strain polymorphisms
lg1
lg2
lg3
lg4
lg5
lg6
lg7
170
85
0
170
85
0
0
283
567
intra−APS4 variants
intra−APS6 variants
inter−strain variants
D
Genotype frequencies of the RAD markers
Genotype frequency
0.8
0.6
0.4
0.2
0.0
LG1
LG2
LG3
LG4
LG5
LG6
LG7
Linkage Group
Homozygous APS4
Homozygous APS6
Heterozygous APS4/APS6


15
10
5
0
log2(count + 1)
L2 Females
L2 Hermaphrodites
Adult Males
Mixed Stages
Chrom
LG1
LG2
LG3
LG4
LG5
LG6
LG7


A
Converted females versus fema
(55)
5
7
Females versus Hermaphrodites
(2422)
32
11
436
160
1943


B
Converted females versus females
log2 FC
More
Converted females versus hermaphrodites
log2 FC
More
More
Females versus hermaphrodites
log2 FC
More
More
genes
LG1
LG2
LG3
LG4
LG5
LG6
LG7


Karyotypes
A. rhodensis
O. tipulae
C. elegans
H. contortus
P. pacificus
1e
2c1 2n1
3a1 3c2
4a2 4c3 4n2
5x
6d
7b
Ia
IIb
IIIc
IVd
Vn12
Xe
Xx
Ia
IIb
IIIc
IVd
Ve
Xn12
Xx
Ia
IIb
IIIc
IVd
Ve
Xn12
Xx
ILe
1Rn12
IIb
IIIc
IVd
Va
Xx
Nigon Elements
NA
a1 a2
NB
b
NC
c1 c2 c3
ND
d
NE
e
NN
n1 n2
NX
x
A. rhodensis
3a1
4a2
7b
2c1
3c2
4c3
6d
1e
4n1
2n2
5x
O. tipulae
Ia
IIb
IIIc
IVd
Xe
Vn12
Xx
C. elegans
Ia
IIb
IIIc
IVd
Ve
Xn12
Xx
H. contortus
1a
2b
3c
4d
5e
Xn12
Xx
P. pacificus
Va
IIb
IIIc
IVd
ILe
1Rn12
Xx